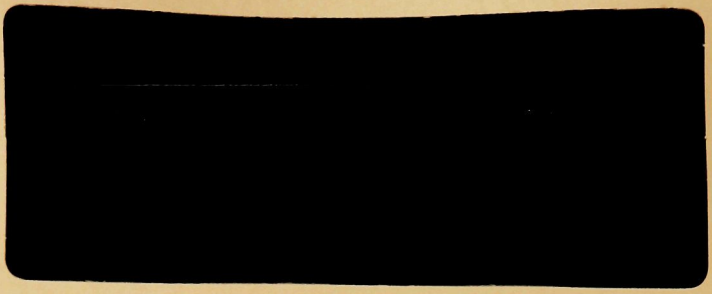
ANL/EES-TM-328

UNCERTAINTY DATA BASE FOR EMISSIONS-ESTIMATION PARAMETERS: INTERIM REPORT

ANL
EES

ARGONNE NATIONAL LABORATORY
Energy and Environmental Systems Division

Operated by
THE UNIVERSITY OF CHICAGO for U. S. DEPARTMENT OF ENERGY
under Contract W-31-109-Eng-38



Argonne National Laboratory, with facilities in the states of Illinois and Idaho, is owned by the United States government, and operated by The University of Chicago under the provisions of a contract with the Department of Energy.

This informal report presents preliminary results of ongoing work or work that is more limited in scope and depth than that described in formal reports issued by the Energy and Environmental Systems Division.

Printed in the United States of America. Available from National Technical Information Service, U. S. Department of Commerce, 5285 Port Royal Road, Springfield, Virginia 22161.

ARGONNE NATIONAL LABORATORY
9700 South Cass Avenue, Argonne, Illinois 60439

ANL/EES-TM-328

UNCERTAINTY DATA BASE FOR EMISSIONS-ESTIMATION PARAMETERS: INTERIM REPORT

by

K.C. Chun

Energy and Environmental Systems Division
Environmental and Resource Assessment Group

March 1987

work sponsored by

U.S. DEPARTMENT OF ENERGY
Office of Fossil Energy
for
National Acid Precipitation Assessment Program
Emissions and Controls Task Group

PREFACE

This report presents the findings of a literature review on uncertainties associated with emissions-estimation parameters and emissions data, analyzes various types of errors and other factors that cause uncertainties in emissions-estimation parameters, and presents the interim results of an effort to compile and develop an uncertainty data base for several of these parameters. Illustrative, preliminary emissions uncertainty estimates are also provided, based on these parameter uncertainties and approximate mathematical procedures. This research is being funded as part of the work of the National Acid Precipitation Assessment Program's Emissions and Controls Task Group by the U.S. Department of Energy (DOE) Office of Fossil Energy. The DOE Project Officer is Edward Trexler.

CONTENTS

FIGURES

TABLES

U.S. DEPARTMENT OF ENERGY PERSPECTIVE

The interim findings contained in this report are part of a modest effort to characterize the uncertainty of the emissions estimates contained in the National Acid Precipitation Assessment Program (NAPAP) emissions data bases. Such data bases vary in resolution from national annual emissions to hourly emissions from particular 20-km grids. Determining uncertainties for the high-resolution portion of the data requires an understanding of the uncertainties of the individual major point sources and of the small urban transportation sources which dominate these small geographic areas. This particular interim report focuses on the range of uncertainties in the major point sources.

While this work builds on the previous uncertainty work of PEDCo and GCA Corporation, the need to provide answers for the high-resolution situation has necessitated a new emphasis on the variances in the parameters used in estimating emissions. Studies of less resolute uncertainties, such as national annual emissions, evolve into studies of parameter biases. This work then will be making a new and significant contribution toward improved understanding of high-resolution uncertainties.

While one considers the potential usefulness of such data, these considerations must be tempered by an understanding of its limitations. Characterization of the uncertainties of the major point source hourly emissions would probably be best accomplished by analyzing emissions test data from representative major point sources; however, sufficient test data are not presently available to support such an approach.

What has been done in the absence of such data has been to mathematically build up such a characterization by characterizing the variances in the emissions-estimation parameters. While we believe that the approach has been generally correct, the application has had limitations. The effort has not had the benefit of data from representatively designed tests; test data were used from many sources. The mathematics employed simplifying assumptions of independence and normality which cannot be fully substantiated by available test data. While we believe that the resultant errors will probably be small and that this is the best attempt yet to acquire such an understanding, we acknowledge the shortcomings and welcome the prospect that improvements will follow and that understanding will be enhanced.

At this particular time, however, when impending NAPAP decisions need to be made in the light of a better understanding of these uncertainties, we believe that it is important to have these interim findings considered.

Edward Trexler, Project Manager
Office of Fossil Energy
U.S. Department of Energy

ACKNOWLEDGMENTS

Many people provided information without which this report could not have been prepared. I am particularly indebted to E. Trexler at the U.S. Department of Energy Office of Fossil Energy for his guidance and assistance throughout the course of this work and for his thorough review of the manuscript. I am also indebted to C. Massey and his staff, Tennessee Valley Authority (TVA), who supplied a considerable amount of emissions and operations data for the TVA system. I would like to acknowledge D. Mobley, C. Mann, T. Lahre, R. Hall, J. Kilgroe, E. Martinez, and C. Sedman, U.S. Environmental Protection Agency (EPA), for helpful discussions and documents; R. Whitfield, Argonne National Laboratory (ANL), for preparing the initial drafts of Sec. 4 and App. C; and G. Bassett, University of Illinois at Chicago, for drafting App. B.

I also wish to thank the following persons for their thorough review of the manuscript and helpful suggestions: D. Hanson, D. Knudson, J. Peerenboom, and D. Streets, ANL; C. Benkovitz, P. Michael, and N. Oden, Brookhaven National Laboratory; J. Ekmann, Pittsburgh Energy Technology Center; J. Langstaff, Alliance Technologies Corp.; J. McDonald, University of Illinois at Chicago; D. Mobley, EPA; H. Ozkaynak, Harvard University, Energy and Environmental Policy Center; B. Price, MITRE Corp.; and L. Stefanski, North Carolina State University.

I am also indebted to P. Chen, ANL, for her careful editing of the report.

UNCERTAINTY DATA BASE FOR EMISSIONS-ESTIMATION PARAMETERS: INTERIM REPORT

by

K.C. Chun

SUMMARY

The National Acid Precipitation Assessment Program (NAPAP) has charged its Task Group on Emissions and Controls (Task Group I) with developing comprehensive and accurate inventories of emissions from man-made and natural sources believed to be important in acid deposition processes. This work involves developing estimates of past, present, and future acid deposition precursor emissions with appropriate geographic, temporal, and source resolution to support the research requirements of NAPAP. Quantifying the degree of uncertainty associated with such emissions estimates is an important NAPAP objective, and is being pursued under several projects. One of these, NAPAP Project B1-19 (titled *Emissions Uncertainties*), which is reported on by this document, is focusing on developing (1) data on the uncertainty due to variability of the emissions-estimation parameters (EEPs), such as emission factors, that are used in computing emissions estimates and (2) methods for calculating emissions uncertainty. Argonne National Laboratory has been compiling uncertainty data for the EEPs and Brookhaven National Laboratory has been working on the methods for calculating emissions uncertainty. The results will contribute to the development of emissions uncertainty data for all of the NAPAP emissions inventories.

The scope of this work is broad, and is intended to provide EEP uncertainty data to various users who calculate emissions estimates in different ways. This interim report provides uncertainty data for the major point source EEPs developed to date. To illustrate the potential applications of these data, this report also presents some preliminary emissions uncertainty estimates that were developed using an approximate methodology. Subsequent Project B1-19 work will focus on uncertainty data for transportation sector area sources. Future work under NAPAP Projects B1-21a and B1-43 will focus on the emissions uncertainties in the monthly state-level emissions inventories from 1975 to the present and in the 1985 detailed emissions inventory.

Emissions estimates at an aggregated level (e.g., a county, state, or national level) or at a point source level may be computed using one of the following two approaches:

1. "Bottom-up" approach, which we call *aggregate-derived*, and

2. "Top-down" approach, which we call *product-derived*.

In the aggregate-derived approach, emissions estimates for individual point sources and smaller area sources are summed to obtain an estimate of total emissions. This approach is presently used in the detailed 1980 and 1985 NAPAP emissions inventories to estimate

state, regional, and national emissions by adding up the emissions from those point sources and county area sources within the area of interest.

In the product-derived approach, emissions estimates are calculated by multiplying relevant EEPs. The product-derived approach is used to make all point source and county-level area source estimates in the detailed 1980 and 1985 NAPAP emissions inventories and to make all point source and state-level area source estimates in the monthly state-level emissions data base. The data provided in this report on EEP uncertainties are only applicable to estimating the uncertainties in product-derived estimates. Brief discussions and examples of determining aggregate-derived uncertainties are provided, but the thrust of this report is to assess uncertainties in product-derived estimates.

In a detailed emissions inventory, estimates of the emissions of a given pollutant from individual point and county-level area sources (representing various source classification code [SCC] categories) may be computed by multiplying the following EEPs:

1. Mean emission factors for the applicable SCC source category,* which contain a term for the mean sulfur or ash content of the fuel (by individual source or county) when applicable,

2. Total activity levels, i.e., total production or throughput (by individual source or county), and

3. Mean emission control system penetration factors (by individual source) when applicable (equal to 1 minus the fractional control or removal efficiency).

The emissions estimates are usually computed for an annual period for each major pollutant. When necessary, such annual emissions are disaggregated by pollutant species, hour, and area size (e.g., 20-km grid) for use in applications such as the Regional Acid Deposition Model (RADM). To obtain such disaggregated data, the annual emissions estimates are multiplied by the following EEPs:

1. Mean pollutant-species speciation factor (by SCC, when applicable),

2. Mean temporal allocation factor‡ (by SCC): quarterly, daily, and hourly, and

3. Area disaggregation factor (by county).

---

*In most, but not all, cases, the emission factor is the arithmetic mean of all available data of acceptable quality.

‡The term *temporal allocation factor* is used throughout this report to refer to the allocation factor used to disaggregate annual activity levels by quarterly, daily, or hourly periods.

In assessing the degree of uncertainty associated with these EEPs, three categories of uncertainty sources can be distinguished: measurement errors, data processing errors (e.g., rounding errors), and representation errors. These errors can be further decomposed into systematic and random components. Systematic components are due to the biases that occur in the measurement process or to nonrepresentative sampling, and random components are errors due to data variability.

Representation errors due to the application of mean values consist of three types:

1. Errors that occur when a mean value based on a limited sample is used to represent the true mean value for the parent population,

2. Errors that occur when a mean value for a given averaging period (e.g., annual) for an individual source that is based on a small number of data points is used to represent the mean value for other averaging periods (e.g., any 1-hr period or group of 1-hr periods, assumed to be randomly selected from the parent population),* and

3. Errors that occur when a mean value from a small sample is used to represent any one or more sources within the same SCC category (assumed to be randomly selected from the present population).‡

The emission factors, pollutant species allocation factors, and temporal allocation factors are cases in which a group mean value (by SCC category) is used to represent the value for an individual source or the mean value for a group of sources or the entire population.

The main purpose of establishing uncertainty ranges for the EEPs is to calculate uncertainty for emissions estimates. In certain situations, such as a field test of major point sources, real-time activity levels and emission factors may be measured at the individual-source level. In such cases, uncertainty estimates based on group mean values for EEPs do not apply. (However, even when individual-source EEPs are measured directly, many causes of variability remain, which may not be detected unless the measurements are made continuously. These causes of variability include changes in load, operating conditions, and fuel characteristics.)

---

*When the sample mean value for an individual source is applied to a specific averaging period for that source, there will remain considerable variability in that averaging period due to such causes as changes in fuel characteristics, load, or operating conditions.

‡When the sample mean value is applied to an individual source for some specific averaging period, there will remain considerable variability in that source for that averaging period due to such causes as changes in individual-source and fuel characteristics, load, or operating conditions.

Uncertainty values are estimated in this study using available data that are based on measured values. Only the uncertainty due to data variability was considered. Systematic error components due to biases that occur in the measurement process or that result from nonrepresentative sampling were not considered.

In order to identify existing EEP uncertainty values, a review was conducted of previous studies on the uncertainties associated with EEPs and emissions data. The most useful of the existing EEP uncertainty values applicable to the NAPAP emissions inventories have been compiled in this report. Development of new uncertainty values from the basic measurement data is in progress. The interim results of this development effort are presented in this report for the benefit of RADM and other researchers. Where basic data are lacking, uncertainty values may eventually be developed by outside experts using interpolation or extrapolation of other EEP uncertainty values.

Interim information on uncertainty is provided in this report for the following EEPs:

1. Annual activity levels for point sources,

2. Coal sulfur content for point sources,

3. Flue-gas desulfurization (FGD) system penetration factors for the electric utility sector,

4. Emission factors for sulfur oxides (SO_x) and nitrogen oxides (NO_x) for fuel combustion point sources, and

5. Temporal allocation factors for the electric utility sector.

For the first EEP listed above, only rounding errors with embedded measurement errors are reported as the measure of uncertainty. For the other four EEPs, which are either individual source- or SCC-group mean values, this report provides actual mean values, along with such associated data as the number of data points on which the mean is based, the variability or spread of the data points (in terms of the coefficient of variation), and the relative extreme values.

The basic data for these EEPs represent a variety of data distribution types. Some distributions appear to be relatively simple and close to the normal distribution, while others are not. To roughly ascertain the magnitude of the uncertainty associated with these parameters in actual applications, relative 95% confidence intervals have been computed, assuming normal distributions, for (1) the parent population mean, (2) the mean value for other averaging periods (i.e., one or more averaging periods of a given length, e.g., 1 hr), and (3) the mean value for one or more sources. These three types of uncertainties correspond to the three types of representation errors discussed earlier. The relative 95% confidence interval is the interval that has a 0.95 probability of containing a given parameter (in this case, the parent population mean [μ] or the mean for one or more [k] sources or averaging periods of the unit length [$\bar{X}_k$]), expressed as a percentage (±%) of the sample mean ($\bar{X}_n$).

The uncertainty values thus compiled and developed, and expressed as relative 95% confidence intervals, are summarized in Table S.1. Typical or mean uncertainty values are given in the table to simplify discussion, but the ranges of values developed in this study for different cases are also given in parentheses for reference. The ranges represent the minimum and maximum values for those cases.

Of the uncertainties due to the three types of representation errors, the largest are those that occur when an individual-source mean value for a given averaging period is used to represent the mean value for other averaging periods (columns 3-5 in Table S.1). The next largest are the uncertainties that occur when an SCC-group mean value is used to represent the mean value for one or more sources (columns 6-8 in Table S.1). The smallest uncertainties are those that occur when a sample mean value is used to represent the mean value for the parent population (column 2 in Table S.1). For some EEPs, the relative 95% confidence intervals listed in Table S.1 extend beyond -100%. Any lower bound that extends beyond -100% is an anomaly caused by assuming normal distribution for data distributions having standard deviations greater than half the mean values. Such lower bounds should be truncated at -100%. In those cases, the reliability of the upper bound is also subject to question.

The uncertainty associated with using a sample mean value to represent the mean value for the parent population (column 2 in Table S.1), expressed as the relative 95% confidence interval, ranges from less than ±1% to ±30%. The utility hourly activity level exhibits the greatest uncertainty of this type (±30%), followed by individual-source NO_x and SO_x emission factors for coal-burning point sources (±30% and ±15%, respectively), then the SCC-group mean NO_x and SO_x emission factors for coal-burning point sources (±15%), utility mean temporal allocation factors (±12% for the quarterly and ±4% for the hourly allocation factors), and the utility FGD system penetration factor (±3%). The coal sulfur content shows a very small uncertainty of this kind (0.5%).

The uncertainty associated with using an individual-source mean value to represent a single hourly averaging period (column 3 in Table S.1) is greatest for the utility hourly activity level, which has a relative 95% confidence interval of -100% to +170% when the lower bound is truncated. The next largest uncertainty is associated with the utility FGD system penetration factor (±100%), followed by the individual-source NO_x and SO_x emission factors for coal-burning point sources (±70% and ±40%, respectively) and the coal sulfur content (±45%). The rounding error associated with the annual activity level shows the smallest uncertainty (typically <5%).

The uncertainty associated with using a sector- or SCC-group mean value to represent a single source (column 6 in Table S.1) ranges from ±30% to ±90% in terms of the relative 95% confidence interval. The utility mean quarterly allocation factor exhibits the greatest uncertainty of this kind (±90%), followed by NO_x and SO_x emission factors for coal-burning point sources (±60% and ±50%, respectively) and the utility mean hourly allocation factor (±30%).

To illustrate the usefulness of these uncertainty estimates for the EEPs, estimates were made of the uncertainties associated with SO_x and NO_x emissions data for a single coal-fired electric utility boiler unit and a group of 100 identical such units, assuming no emission control devices. The estimates were based on the mean values for

TABLE S.1 Summary of Uncertainty Values Identified or Developed in This Report for Selected EEPs[a] (95% confidence intervals, expressed as ±% of the EEP value)

EEPs by Category (col. 1)	Uncertainty Due to Representation Errors: When the EEP Represents the Parent Population Mean (col. 2)	When the EEP Represents the Mean for One Averaging Period: Hour (col. 3)	Day (col. 4)	Quarter (col. 5)	When the EEP Represents the Mean for k Sources[b]: k = 1 (col. 6)	k = 10 (col. 7)	k = 100 (col. 8)	Uncertainty Due to Rounding Errors (col. 9)
Individual-source values: annual activity level								<5 (0.1-50)
Individual-source mean values								
Coal sulfur content[c]	0.5 (0.3-0.8)	45 (20-80)	15 (10-20)	2 (1-4)				
FGD system penetration factor[c]	3 (2-5)	100 (60-130)	85 (40-120)					
Emission factor								
SO_x from coal combustion	15 (5-30)	40 (15-70)						
NO_x from coal combustion	30 (25-50)	70 (50-120)						
Hourly activity level[c]	30 (0.5-200)	170 (3-1600)						
SCC-group mean values								
Emission factor								
SO_x from coal combustion	15 (5-20)				50 (40-70)	20 (15-30)	15 (10-25)	
NO_x from coal combustion	15 (5-600)				60 (40-1100)	25 (15-400)	20 (10-300)	
Temporal allocation factor[c]								
Quarterly	12 (10-15)				90 (60-130)	30 (20-45)	15 (10-20)	
Hourly	4 (1-50)				30 (10-380)	10 (3-130)	5 (2-60)	

[a]Each value listed is the mean for the range of values given in parentheses below it. Confidence intervals that extend beyond -100% should be truncated at -100%.

[b]Randomly selected from the parent population.

[c]For the electric utility sector.

the uncertainty associated with the pertinent EEPs. Approximate mathematical procedures and several simplifying assumptions were used. The bounds for the relative 95% confidence intervals thus estimated for SO_x and NO_x emissions over various time periods (i.e., annual, quarterly, a mean specific hourly period, and an individual specific hour) are listed in Table S.2.

For a single electric utility unit, the relative 95% confidence intervals for SO_x emissions (column 2 in Table S.2) are quite large, ranging from a minimum interval of about ±40% for an annual period to a maximum interval of -100% to about +230% for an individual hour. The relative 95% confidence intervals for NO_x emissions are comparable, ranging from a minimum interval of about ±70% for an annual period to a maximum interval of -100% to about +230% for an individual hour.

The emissions uncertainty estimates become smaller for a group of units than for a single unit. For a group of 100 units, the relative 95% confidence intervals (column 3, product-derived approach) range from ±7% for an annual period to ±25% for an individual hourly period for SO_x emissions and from ±15% for an annual period to ±28% for an individual hourly period for NO_x emissions. Uncertainty intervals for the 100-unit case based on the aggregate-derived approach are smaller than those based on the product-derived approach in all cases. The reason for the difference is that the aggregate-derived approach assumes that the values for the SCC-group mean EEPs have been estimated for each individual unit, and therefore fails to treat the uncertainty due to an insufficient amount of measurement data used in developing mean EEP values, which leads to variability in knowing the true population mean.

These illustrative, preliminary uncertainty estimates are rough approximations based on a number of simplifying assumptions and an approximate methodology. In addition, these estimates are applicable to average or typical situations only, because they are based on average or typical values for the magnitude of uncertainty associated with the pertinent EEPs. Depending on the situation, e.g., whether the units under consideration are baseload or intermediate-load units, or whether the time period of concern falls on a weekday or weekend, the emissions uncertainty estimates could be larger or smaller.

While these estimates are derived for the electric utility sector, they might suggest the order of magnitude of the uncertainty associated with SO_x and NO_x emissions from other point source combustion processes. Combustion point sources currently account for about 80% of the SO_x emissions and 50% of NO_x emissions in the United States.

TABLE S.2 Example of Uncertainty Estimates for SO_x and NO_x Emissions from One or a Group of 100 Identical Coal-Fired Electric Utility Boiler Units[a]

Pollutant, Period (col. 1)	Relative 95% Confidence Interval[b] (±%)		
		Emissions from 100 Units	
	Single-Unit Emissions (col. 2)	Product-Derived Approach (col. 3)	Aggregate-Derived Approach[c] (col. 4)
SO_x			
Year	41	7	4
Quarter	101[d]	17	10
Hour			
Mean	106[d]	17	11
Individual	228[d]	25	23
NO_x			
Year	70	15	7
Quarter	118[d]	21	12
Hour			
Mean	123[d]	22	12
Individual	234[d]	28	23

[a]No emission control is assumed.

[b]Expressed as ±% of the emissions estimate for each period.

[c]This approach neglects the uncertainty due to an insufficient amount of measurement data used in developing mean EEP values.

[d]Confidence intervals that extend beyond -100% should be truncated at -100%.

1 INTRODUCTION

The National Acid Precipitation Assessment Program (NAPAP) has charged its Task Group on Emissions and Controls (Task Group I) with developing comprehensive and accurate inventories of emissions from man-made and natural sources believed to be important in acid deposition processes. This work involves developing estimates of past, present, and future acid deposition precursor emissions with appropriate geographic, temporal, and source resolution to support the research requirements of NAPAP. Quantifying the degree of uncertainty associated with such emissions estimates is an important NAPAP objective that is being pursued under several projects. One of these, NAPAP Project B1-19 (titled *Emissions Uncertainties*), which is reported on by this document, is focusing on developing (1) data on the uncertainty due to variability of the emissions-estimation parameters (EEPs), such as emission factors, that are used in computing emissions estimates and (2) methods for calculating emissions uncertainty. Argonne National Laboratory (ANL) has been compiling EEP uncertainty data and Brookhaven National Laboratory (BNL) has been working on the methods for calculating emissions uncertainty.

The scope of this work is broad, and is intended to provide EEP uncertainty data to various users who calculate emissions estimates in different ways. This interim report provides uncertainty data for the major point source EEPs developed to date. To illustrate the potential applications of these data, this report also presents some preliminary emissions uncertainty estimates that were developed using an approximate methodology. Subsequent B1-19 work will focus on uncertainty data for transportation sector area sources. Future work under NAPAP Projects B1-21a and B1-43 will focus on the emissions uncertainties in the monthly state-level emissions inventories from 1975 to the present and in the 1985 detailed emissions inventory.

Emissions estimates at an aggregated level (e.g., a county, state, or national level) or at a point source level may be computed using one of the following two approaches:

1. "Bottom-up" approach, which we call *aggregate-derived*, and

2. "Top-down" approach, which we call *product-derived*.

In the aggregate-derived approach, emissions estimates for individual point sources and smaller area sources are summed to obtain an estimate of total emissions. This approach is presently used in the detailed 1980 and 1985 NAPAP emissions inventories to estimate state, regional, and national emissions by adding up the emissions from those point sources and county area sources within the area of interest.

In the product-derived approach, emissions estimates are calculated by multiplying relevant EEPs. The product-derived approach is used to make all point source and county-level area source estimates in the detailed 1980 and 1985 NAPAP emissions inventories and to make all point source and state-level area source estimates in the monthly state-level emissions data base. The data provided in this report on EEP

uncertainties are only applicable to estimating the uncertainties in product-derived estimates. Brief discussions and examples of determining aggregate-derived uncertainties are provided, but the thrust of this report is to assess uncertainties in product-derived estimates.

In a detailed emissions inventory, annual emissions are first estimated for individual point sources and county-level area sources representing various source classification code (SCC) categories for a number of pollutants. Such annual emissions (E_y) are computed by multiplying the following EEPs:

$$E_y = (EF)(A)(PF)^* \tag{1.1}$$

where:

EF = mean emission factor, by SCC group (this EEP contains a term for the annual mean fuel ash or sulfur content by individual source or county, when applicable),

A = total annual activity level (i.e., production or throughput), by individual source or county, and

PF = annual mean control system penetration factor, by individual source (this term equals 1 minus the fractional removal efficiency).

The annual emissions data thus obtained can be disaggregated by pollutant species, shorter time periods, and smaller area size for area sources (i.e., 20-km grid areas) for use in such applications as the Regional Acid Deposition Model (RADM), which is currently being developed by the U.S. Environmental Protection Agency (EPA). Temporally allocated emissions (E_t) are obtained by multiplying the annual emissions data from Eq. 1.1 by the following EEPs, as applicable:

$$E_t = (E_y)(SF)^*(TF)(AF) \tag{1.2}$$

where:

E_t = emissions for shorter periods than a year (e.g., hourly), by individual source or grid and by pollutant species,

SF = mean pollutant species speciation factor, by SCC group,

*This term applies when appropriate, but otherwise equals 1.

TF = mean temporal allocation factor,* and

AF = area disaggregation factor, by county.

One of the objectives of NAPAP Project B1-19 is to develop uncertainty values for the following EEPs:

1. *Emission factors* for each possible combination of 29 point and area source categories and 6 pollutant species: sulfur dioxide (SO_2), sulfate, nitrogen oxides (NO_x), total volatile organic compounds, total particulate matter, and ammonia;

2. *Annual activity levels* for five major sectors: electric utilities, industrial combustion, industrial processes, transportation, and the residential/commercial sector;

3. *Sulfur and ash contents of fuels*;

4. *Penetration factors* for utility and industrial flue gas desulfurization (FGD) systems and particulate control systems;

5. *Temporal allocation factors* for SO_2 and all other pollutants from each of the five above-mentioned sectors and for the following time categories: calendar quarter, month, weekday hour, and weekend hour; and

6. *Spatial disaggregation factors* for major area source categories.

In assessing the degree of uncertainty associated with these EEPs, three categories of uncertainty sources can be distinguished: measurement errors, data processing errors (e.g., rounding errors), and representation errors. These errors can be further decomposed into systematic and random components. Systematic components are due to the biases that occur in the measurement process or to nonrepresentative sampling, and random components are errors due to data variability.

Representation errors due to the use of mean values consist of three types:

1. Errors that occur when a mean value based on a limited sample is used to represent the true mean value for the parent population,

2. Errors that occur when a mean value for one averaging period (e.g., a year) and one individual source that is based on a small number of data points is used to represent the mean value for

*The term *temporal allocation factor* is used throughout this report to refer to the allocation factor used to disaggregate annual activity levels by quarterly, daily, or hourly periods.

other averaging periods (e.g., any 1-hr period or group of 1-hr periods, assumed to be randomly selected from the parent population),* and

3. Errors that occur when a mean value from a small sample is used to represent any one or more sources within the same SCC category (assumed to be randomly selected from the parent population).‡

The emission factors, pollutant species allocation factors, and temporal allocation factors are cases in which a group mean value (by SCC category) is used to represent the value for an individual source or the mean value for a group of sources or the entire population.

The main purpose of establishing uncertainty ranges for the EEPs is to calculate uncertainty for emissions estimates. In certain situations, such as a field test of major point sources, real-time activity levels and emission factors may be measured at the individual-source level. In such cases, uncertainty estimates based on group mean values for EEPs do not apply. (However, even when individual-source EEPs are measured directly, many causes of variability remain, which may not be detected unless the measurements are made continuously. These causes of variability include changes in load, operating conditions, and fuel characteristics.)

The approach taken to acquire the EEP uncertainty values is to:

1. Review previous studies of EEPs and emissions data uncertainties and compile existing applicable uncertainty values for EEPs,

2. Develop additional uncertainty values for EEPs using available basic data, when applicable values are not available, and

3. Use experts in the field to interpret and confirm the EEP uncertainty values compiled and developed, and to interpolate or extrapolate these values to similar source and pollutant species categories when basic data are lacking.

---

*When the sample mean value for an individual source is applied to a specific averaging period for that source, there will remain considerable variability in that averaging period due to such causes as changes in fuel characteristics, load, or operating conditions.

‡When the sample mean value is applied to an individual source for some specific averaging period, there will remain considerable variability in that source for that averaging period due to such causes as changes in individual-source and fuel characteristics, load, or operating conditions.

Section 2 of this interim report analyzes the characteristics of errors that cause the EEP uncertainties and outlines statistical measures for expressing the degree of these uncertainties. Section 3 presents the uncertainty values identified or developed to date for selected EEPs (primarily for major combustion point sources). Illustrative, preliminary emissions uncertainty estimates are described in Sec. 4. A summary is provided in Sec. 5. Appendix A summarizes the findings of the literature review, App. B presents the statistical formulas for describing EEP uncertainties and procedures for their aggregation, and App. C describes the mathematical procedures and assumptions used in estimating emissions uncertainties. Appendix D provides the basic data used in this study.

2 SOURCES AND CHARACTERISTICS OF EEP UNCERTAINTIES

The hourly rate of emissions of a pollutant species from an individual source (or 20-km grid) will be computed using Eqs. 1.1 and 1.2 for NAPAP research purposes. For analyzing the characteristics of the sources of uncertainties associated with the EEPs in those equations, it is useful to classify these parameters into three categories, as shown in Table 2.1. Examples of these three categories are, in the order shown in Table 2.1, (1) the annual amount of coal consumption at a point source, (2) the annual mean coal sulfur content at a point source, and (3) the emission factor for an SCC category.

Accurate values for EEPs cannot be obtained for all emission sources in the nation due to cost constraints on data gathering and the inherent variability in measurement techniques. To approximate the values, two approaches are possible. One is to use mean values based on the limited available amounts of measured data. The other approach, when measured data are not available, is to use estimates based on engineering analysis or judgment. Both approaches result in errors that cause EEP uncertainties. Although errors associated with the latter approach (estimation errors) are difficult to analyze, those associated with the former can be classified into several categories for analysis: (1) measurement errors, (2) data processing errors, and (3) representation errors.

TABLE 2.1 Categories of EEPs

Category	EEP
Individual-source or individual county values	annual activity level spatial disaggregation factor
Individual-source mean values	sulfur or ash content of fuel emission control equipment penetration factor hourly activity level
SCC-group or individual-county mean values	
SCC group[a]	emission factor pollutant speciation factor temporal allocation factor
County	sulfur or ash content of fuel for area sources

[a]Such groups can consist of one SCC or multiple SCCs.

Measurement errors occur due to to the variability in measurement techniques. Data processing errors include derivation, rounding, and data transfer errors. Derivation errors arise when a measured value is converted to an EEP value through certain mathematical manipulations, sometimes involving the use of the mean values of some other variables. Rounding errors result when data are rounded up or down according to industry conventions. Data transfer errors occur when mistakes are committed during such processes as coding. Improved quality assurance procedures adopted for the development of the NAPAP emissions inventories are expected to reduce data transfer errors.

Measurement errors and data processing errors are associated with the individual and mean values of all three categories of EEPs. Representation errors occur when the mean EEP values are used in calculating annual emissions or temporally and spatially disaggregated emissions. For example, a representation error occurs when the annual mean fuel sulfur content for a point source is used in calculating SO_2 emissions at that point source for a single, specific hourly period.

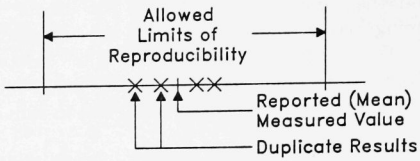
These errors can be further decomposed into systematic and random components. Systematic components are due to biases that occur in the measurement process or as a result of nonrepresentative sampling, and random components are due to data variability. The decomposition of total error into these components is illustrated in Fig. B.2. Systematic errors are often not easily quantified, but can be minimized by designing and collecting representative samples of data and/or correcting the data for measurement biases. However, random errors due to data variability can be analyzed statistically for quantification.

The magnitude of uncertainty associated with measurement and rounding errors can be quantitatively expressed by use of relative error bounds (error bounds divided by the parameter value), which are shown in Fig. 2.1. For example, in this report, for measurement errors in the sulfur or ash contents of individual samples of solid and liquid fuels, the limits of reproducibility allowed by the American Society of Mechanical Engineers (ASME) test codes[1,2] are used as the error bounds. For rounding errors, the maximum range of the values rounded to a reported value is used as the error bound for the reported value.[3] As shown in Fig. 2.1a, all duplicate measurements should fall within the allowed limits of reproducibility, and all data that are rounded up or down should fall within the maximum range of the values to be rounded. In other words, the error bounds contain all of the individual sulfur or ash content measurement values before rounding. These intervals can be considered to be approximately equal to the 95% confidence interval commonly used in expressing uncertainties associated with estimated values such as emissions data. (A 95% confidence interval is an interval that has a 0.95 probability of containing a given value.)

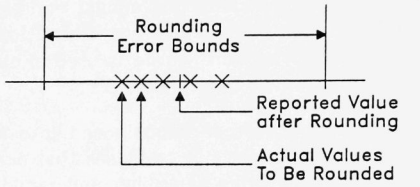
Statistical formulas to express the magnitude of uncertainty due to representation errors are not as simple as those for measurement or rounding errors. In App. B, confidence intervals are derived for sample mean values to represent parent population means, as well as finite populations other than the sample population. As an example, the 95% confidence intervals for the population mean and the mean value for a finite population (when size $k = 1$) are illustrated in Fig. 2.1b for a normally distributed population with a sample size of $n = 30$.

(a) Error Bounds (100% confidence intervals)

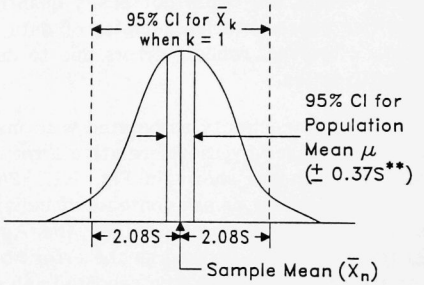
(i) for measurement error



(ii) for rounding error



(b) 95% Confidence Intervals (CI) for Mean Values *



* Illustrated for a normal distribution with n = 30

** S = Standard deviation

FIGURE 2.1 Error Bounds and 95% Confidence Intervals

3 COMPILATION AND DEVELOPMENT OF EEP UNCERTAINTY VALUES

This section presents, for selected EEPs, uncertainty values that have been either newly developed or identified from the literature as directly relevant to the NAPAP emissions inventories. The following EEPS are covered:

1. Annual activity level for point sources,
2. Coal sulfur content for point sources,
3. FGD system penetration factors for the electric utility sector,
4. Emission factors for sulfur oxides (SO_x) and NO_x for fuel combustion point sources, and
5. Temporal allocation factors for the electric utility sector.

3.1 DESCRIPTION OF THE DATA USED

In the past, uncertainty values for EEPs have been based primarily on the following two methods: (1) analysis of data (measured, reported, or survey data) and/or (2) engineering judgment or engineering analysis by an individual or panel. Since the latter method, if not carefully designed and applied, sometimes lacks objectivity and/or credibility, only those uncertainty values derived from measured or reported data were considered for data compilation in this study.

Uncertainty estimates based on measured data currently available in the literature are largely limited to criteria pollutant/major process combinations developed by PEDCo.[3] No such estimates are available for emission factors for pollutants such as ammonia, for which emission factors have just recently been developed, and few estimates are available for species, temporal, and spatial allocation factors. In addition, most of the available uncertainty values for emission factors are not suitable for the NAPAP emissions inventory uncertainty assessment, because they were developed for the EPA emission factors (referred to as AP-42 emission factors) published in 1973,[4] not the revised AP-42 emission factors (published in 1985)[5] to be used in the development of the NAPAP emissions inventories.

The presentation of uncertainty data is relatively simple in the cases of measurement errors or rounding errors. However, presentation of the data variability that causes representation errors is somewhat more involved. For this purpose, a largely self-explanatory data form (Fig. 3.1) has been developed. The top half of the form is used for presenting the data source, mean value (X), extreme values, sample standard deviation (S), and number of observations or data points (n). In the case of emission factors, the AP-42 emission factors compiled by EPA are listed first for comparison purposes. The bottom half of the form is used for presenting coefficient of variation (CV) values and relative extreme values with respect to the mean value. The definitions

Min. 1.0S 1.0S Max. $\bar{X}$

Emissions-Estimation Parameter Variability Data Base

Low REV 1.0CV 1.0CV High REV 0

Date:

Parameter		
SCC	Number	
	Description	

Data Source	Mean ($\bar{X}$), Standard Deviation (S), and Max. and Min. Values	Sample Number

Data Source	Coefficient of Variation (CV), % and Relative Extreme Value (REV), %	

Comments:

FIGURE 3.1 Data Form for the EEP Variability Data Base

of the statistical parameters follow standard conventions, as described in App. B.* Completed data forms are included in App. D as background information.

The basic data for these EEPs represent a variety of data distribution types. Some distributions appear to be relatively simple and close to the normal distribution, while others are not. For example, the frequency distribution of hourly net generation data for a Tennessee Valley Authority (TVA) electric generating unit, Allen Unit 2, during the first quarter of 1985,[6] is very much different from normal (see Fig. 3.2).

To roughly ascertain the magnitude of the uncertainty associated with these EEPs in actual applications, relative 95% confidence intervals were computed, assuming normal distributions, for (1) the parent population mean, (2) the mean value for one or more averaging periods of a given length (e.g., an hour), and (3) the mean value for one or more sources. These three types of uncertainties correspond to the three types of representation errors discussed in Sec. 1. The relative 95% confidence interval is the interval that has a 0.95 probability of containing a given parameter (in these cases, the parent population mean [μ] or the mean for one or more sources or averaging periods of a given length [$\bar{X}_n$]), expressed as a percentage (±%) of the sample mean ($\bar{X}_n$). For some EEPs, the relative 95% confidence intervals extend beyond -100%, which is an anomaly

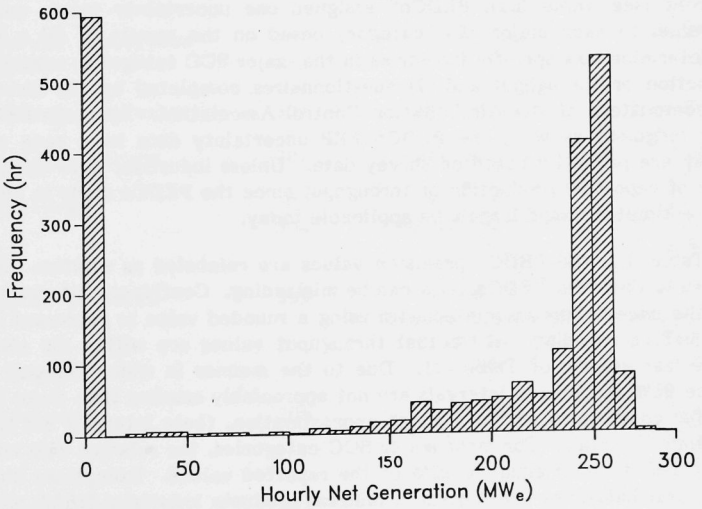


FIGURE 3.2 Example of Nonnormal Data Distribution: Hourly Net Generation Data for a TVA Generating Unit (first quarter, 1985)

*A relative extreme value is defined as follows: [(maximum or minimum value - mean value)/mean value] x 100%.

caused by assuming normal distribution for data distributions having standard deviations that are greater than half the mean values. All such values should be truncated at -100%. In those cases, the reliability of the upper bounds is also subject to question.

3.2 ANNUAL ACTIVITY LEVELS

The annual activity levels (fuel or materials throughput) for point sources to be used in developing the NAPAP emissions inventories are to be obtained from company records when such data are available.[7] Since the annual activity level is a numerical value for an individual source, there are no representation errors due to data variability. The major source of uncertainty would be data processing errors, including rounding errors that result from industry-specific conventions in reporting production or throughput data. Measurement errors are embedded in the rounding error. Errors caused by mistakes in data handling processes may be large, but cannot be easily quantified. Therefore, only rounding errors are considered for this EEP. Additional uncertainty could be introduced if surrogate values, instead of the actual production or throughput values directly linked to the emission factors, are used. However, such additional uncertainty is not considered in this report.

For annual production or throughput data based on company records for individual sources (see Table 3.1), PEDCo[3] assigned one uncertainty value, calling it a *precision value*, to each major SCC category based on the results of (1) a literature search to determine how specific industries in the major SCC categories normally report their production or throughput and (2) questionnaires completed by the directors and executive committees of the Air Pollution Control Association. Although PEDCo used engineering judgment as well, the PEDCo EEP uncertainty data have been used here because they are primarily based on survey data. Unless industries have changed their conventions of reporting production or throughput since the PEDCo study in 1974, these uncertainty estimates should largely be applicable today.

In Table 3.1, the PEDCo precision values are relabeled as relative 95% confidence intervals, since the PEDCo term can be misleading. Confidence intervals are used to express the uncertainty associated with using a rounded value to represent an actual value, i.e., before rounding. All actual throughput values are within the ±% interval listed in the last column of Table 3.1. Due to the manner in which industry data are reported, the 95% confidence intervals are not appreciably smaller than these intervals. Therefore, for consistency and as a rough approximation, these intervals are treated as 95% confidence intervals. For most major SCC categories, the relative 95% confidence intervals are less than or equal to ±2% of the reported values. Exceptions include the food/agricultural industries (±10%); small mineral products industries (±5%); wood products industry (±1% to ±5%); metal fabrication, leather products, and textile products industries and other or not classified categories (±5%); and solid waste disposal operations (±10% to ±50%). As indicated before, the uncertainty values in Table 3.1 account only for rounding errors. Therefore, these values must be considered as the lowest limit of uncertainty in the annual throughput for a single source.

TABLE 3.1 Uncertainty Values for Annual Activity Levels[a]

Industry	SCC (first 3 digits)	Amount of Throughput (any unit of measure)	Relative 95% Confidence Interval (±%)[b]
Electricity generation	101	>200,000	0.1
		≥10,000-200,000	20,000/X[c]
		<10,000	0.1
Industrial boilers	102	>200,000	0.1
		>50,000-200,000	50,000/X[c]
		≥1,000-50,000	2.0
		<1,000	0.1
Commercial/ institutional boilers	103	>200,000	0.1
		>50,000-200,000	50,000/X[c]
		≥1,000-50,000	2.0
		<1,000	0.1
Internal combustion engine	201-288	All values	0.2
Chemical manufacture	301	All values	1.0
Food/agriculture	302	All values	10.0
Metallurgy	303	>500,000	0.1
	304	≥50,000-500,000	50,000/X[c]
		<50,000	1.0
Mineral products	305	>1,000,000	0.1
		>50,000-1,000,000	100,000/X[c]
		≥10,000-50,000	2.0
		<10,000	5.0
Petroleum	306	>200,000	0.1
		≥10,000-200,000	20,000/X[c]
		<10,000	0.1
Wood products	307	≥10,000	1.0
		<10,000	5.0
Metal fabrication	309	All values	5.0
Leather products	320	All values	5.0
Textile products	330	All values	5.0

TABLE 3.1 (Cont'd)

Industry	SCC (first 3 digits)	Amount of Throughput (any unit of measure)	Relative 95% Confidence Interval (±%)[b]
In-process fuel	390	>200,000	0.1
		>50,000-200,000	50,000/X[c]
		≥1,000-50,000	2.0
		<1,000	0.1
Other/not classified	399	All values	5.0
Point source evaporation	401-490	All values	1.0
Solid waste disposal	501-503	>10,000	10.0
		≥1,000-10,000	20.0
		<1,000	50.0

[a]Due to industry reporting conventions, i.e., rounding errors.

[b]±% of the throughput value for one source, derived from PEDCo's fractional values.

[c]Formula to calculate the relative 95% confidence interval, where X = the throughput value.

3.3 SULFUR AND ASH CONTENTS OF FUELS

The individual-source fuel sulfur and ash contents to be used in developing the NAPAP emissions inventories are to be annual weighted mean values.[7] This section considers the uncertainties associated with (1) measurements of fuel sulfur and ash contents and (2) the inherent variability in fuel sulfur and ash contents.

Uncertainty values for fuel sulfur and ash content data were developed by PEDCo[3] considering measurement errors only, based on the specifications for sulfur and ash content analyses described in the ASME test codes.[1,2] The uncertainty values listed in Table 3.2 were derived by dividing the allowed range of reproducibility (for coal, the ± permissible differences when tested in different laboratories and, for oil, the ± deviation from the mean of measurements by different operators or apparatus) by the reported sulfur or ash content value. Although PEDCo referred to these uncertainty values as precision values, they are actually the intervals that contain all measured values obtained by different laboratories for coal and by different operators or apparatus for oil. For the same reasons given for the annual activity levels in Sec. 3.2, these intervals are reported as relative 95% confidence intervals in Table 3.2. The relative 95%

TABLE 3.2 Uncertainty Values for Fuel Sulfur and Ash Content Data[a]

Impurity, Fuel Type (col. 1)	Reported Sulfur or Ash Content (%) (col. 2)	Formula to Calculate the Relative 95% Confidence Interval[b] (col. 3)	Examples: Sulfur or Ash Content (%) (col. 4)	Examples: Relative 95% Confidence Interval (±%)[c] (col. 5)
Sulfur				
Solid fuels	<2.0	10/X	1.0	10
	≥2.0	20/X	4.0	5.0
Distillate and residual oil	≤0.5	3/X	0.3	10
	>0.5-1.0	4/X	0.75	5.3
	>1.0-2.0	5/X	1.5	3.3
	>2.0-3.0	7/X	2.5	2.8
	>3.0-4.0	9/X	3.5	2.6
	>4.0-5.0	12/X	4.5	2.7
	>5.0	14/X	6.0	2.3
Ash: solid fuels only	≤12.0	50/X	10	5.0
	>12.0	100/X	15	6.7

[a]Due to measurement errors.

[b]Where X is the reported sulfur or ash content value (see col. 2).

[c]±% of the sulfur or ash content value.

confidence intervals range from about ±2% to ±10% for the sulfur and ash content of solid and liquid fuels.

The uncertainty values listed in Table 3.2 account only for the measurement error for individual samples. Additional uncertainty may arise due to the inherent variability of the sulfur and ash content within a fuel supply. For example, the sulfur content of a given coal seam is not uniform, but varies throughout the seam. Therefore, the average sulfur content of a coal lot is closely related to that of previous lots (i.e., it is autocorrelated), but varies as a function of lot size. For a coal-burning boiler, lot size is determined as the product of the coal consumption rate (a function of boiler size) and the duration of firing (averaging time).

For assessing the uncertainty associated with coal sulfur content due to its inherent variability, long-term data for short averaging periods (e.g., hourly) are needed. However, such data are rarely available. Furthermore, the annual weighted-average

values to be used for the NAPAP emissions inventories, even if available, would not provide such data. Therefore, it is not possible to conduct an uncertainty assessment based on reported, individual-source sulfur content data. However, since the effects of averaging time and autocorrelation can be approximated by a first-order autoregressive model[8] (assuming that the sulfur content of a coal lot is related only to that of its previous lot), such a model can be used to estimate the CV of sulfur content for a typical coal for various boiler sizes and averaging times. This analysis assumes that the coal lots come from the same source and that the long-term average sulfur content does not change.

Table 3.3 shows the sulfur content variability in raw and cleaned coal in terms of CV based on a first-order autoregressive model and representative values of CV for reference coal lots and autocorrelation coefficients.[8] The CV values decrease as the boiler size increases and as the averaging period lengthens. Limitations of the current coal sulfur analysis (CSA) process make it impossible to separate errors related to the CSA process from those associated with the inherent variability in coal properties. Since the representative values of CV for reference coal lots were derived using CSA data, the errors related to the CSA process are embedded in the CV values listed in Table 3.3.

Although this particular autocorrelated variable is not quite normally distributed, it can be represented fairly well by a normal distribution.[9] It was also assumed, in the derivation of the model-based CV values, that the number of samples for each averaging period is equal to the number of such averaging periods in a year. Based on these two assumptions, one can use Eq. B.4 to estimate 95% confidence intervals for the population mean (a measure of uncertainty when the sample mean $\bar{X}$ is used to represent the parent population mean μ) and Eq. B.5 to estimate 95% confidence intervals for the mean when it is used to represent a single averaging period (a measure of uncertainty when the sample mean is used to represent a single sample value, e.g., the value for 1 hr).

The confidence intervals so computed are also listed in Table 3.3. The relative 95% confidence intervals for the population mean are less than or equal to about ±1%, reflecting the effects of the large number of samples assumed for short-term averaging periods. These intervals do not have much relevance to actual situations, since CSA data may be limited. The relative 95% confidence intervals for the mean when it is applied to a single averaging period are rather large for the hourly averaging period (e.g., about ±80% for a 100-MWe boiler burning raw coal), suggesting that a large degree of uncertainty may be introduced (due to the inherent variability of coal properties) when annual emissions are temporally allocated to generate detailed emissions data for applications such as RADM. It should be noted that the relative 95% confidence intervals presented in Table 3.3 are applicable only to the ideal case postulated, using coal from a single mine sequentially as the coal is mined and delivered. Many electric utility companies acquire coal from multiple sources and blend coals from different sources to meet required limits on SO_2 emissions.

3.4 CONTROL EQUIPMENT PENETRATION FACTORS

The emission control equipment efficiency data to be used for the NAPAP emissions inventories are to be measured values at individual sources (or design

TABLE 3.3 Variability and Uncertainty Values for the Sulfur Content of Raw and Cleaned Coal

Averaging Period, Coal Type	Boiler Size (MWe)[a]	No. of Samples[c]	CV[b] (%)	Relative 95% Confidence Interval (±%)	
				For the Parent Population Mean[d]	For the Mean for One Averaging Period[e]
Hourly					
Raw	100	8760	40	0.84	78
	500	8760	21	0.44	41
Cleaned	100	8760	19	0.40	37
	500	8760	12	0.25	24
Daily					
Raw	100	365	12	1.2	24
	500	365	5.7	0.58	11
Cleaned	100	365	8.6	0.88	17
	500	365	4.5	0.46	8.8
Quarterly					
Raw	100	4	1.4	2.2	5.0
	500	4	0.60	1.0	2.1
Cleaned	100	4	1.1	1.8	3.9
	500	4	0.49	0.78	1.7
Yearly					
Raw	100	1	0.68	-	-
	500	1	0.30	-	-
Cleaned	100	1	0.54	-	-
	500	1	0.24	-	-

[a]The coal consumption rate is assumed to be 0.4 ton/hr/MWe.

[b]Based on a first-order autoregressive model, using (1) autocorrelation coefficients of 0.25 and 0.50 for raw and cleaned coal, respectively, and (2) CVs of sulfur content for a 240-ton reference lot of 0.20 and 0.12, respectively, for raw and cleaned coal.[8]

[c]Since CV values are generated by use of a model, the numbers of samples are taken as those available within a 1-yr period.

[d]Based on Eq. B.4 (see Sec. B.2.1).

[e]Based on Eq. B.5 (see Sec. B.2.2).

efficiency values, if measured efficiencies are not available), with necessary adjustments to account for control equipment downtime. For the 1985 inventory, these data are to be expressed in weight percent to the nearest 0.1%, and should represent an annual average of the control efficiency actually achieved.[7] Thus, they are subject to all of the types of errors discussed in Sec. 2.1.

PEDCo[3] developed what it called precision values for the control equipment efficiency data based on how precisely the efficiency data were reported, which accounted only for rounding errors. However, the PEDCo values are no longer useful, at least for the 1985 NAPAP emissions inventory uncertainty assessment because, as indicated above, all control equipment efficiency data are to be reported to the nearest 0.1%. This required degree of detail seems to be far greater than may be warranted when one considers the large degree of uncertainty caused by some other sources of errors.

The uncertainties in control equipment efficiency data caused by measurement and derivation errors can be substantial. To assess these uncertainties for any actual data set, a detailed review of background information on measurement techniques and procedures for deriving efficiency data from raw data is needed.

Actual efficiency measurements for individual-source emission control equipment are subject to representation errors due to the inherent variability in the control efficiency, which results from changes in emission source and control equipment operating conditions and from changes in ambient conditions. If a design efficiency value or a manufacturer's guaranteed efficiency value is used, then an estimation error results. To assess the potential levels of these uncertainties, an analysis was performed of wet scrubber SO_2 removal efficiency data collected in an EPA performance testing program for electric utility FGD systems.[10] The testing program involved two dual alkali and two limestone-assisted wet scrubber systems, which were considered to be state-of-the-art systems representative of the majority of scrubbers installed between 1975 and 1980. Only the data obtained while the FGD systems were in operation were used in the analysis. During periods of temporary system failure, emission rates could increase drastically. Therefore, the uncertainty values estimated below should be considered as the lowest limits of the potential uncertainty associated with the utility FGD system penetration factor (which is equal to 1 minus the fractional control efficiency).

The mean values, number of data points on which each mean value was based, variabilities in terms of CV, and relative extreme values for the hourly and daily SO_2 penetration factors are listed in Table 3.4 for five data sets covering four FGD systems. In general, the variability in the penetration factor for each data set is substantial, with CV values ranging from about 30% to 70% for the hourly averaging period (with a mean value of about 50%) and from about 20% to 60% for the daily averaging period (with a mean value of about 40%).

The relative 95% confidence intervals shown in Table 3.4 were computed assuming normal distribution. The relative 95% confidence intervals for the population mean penetration factors range from ±2% to ±5% and from ±6% to ±22% for the hourly and daily averaging periods, respectively. This indicates that the uncertainties associated with using the sample mean penetration factor to represent the population

TABLE 3.4 Variability and Uncertainty Values for FGD System Penetration Factors[a]

Averaging Period, FGD System	Penetration Factor[a] (%)				No. of Samples	CV (%)	Relative Extreme Values (%)		Relative 95% Confidence Interval (±%)[b]	
	Design	Guaranteed	Measured Mean						For the Parent Population Mean[c]	For the Mean for One Averaging Period[d]
			Hourly	Daily			Low	High		
Hourly										
Dual alkali										
FMC Corp.[e]	12	20	12.7	--	1299	30	-84	+109	2	59
CEA/ADL[f]	5	10	7.2	--	2483	43	-82	+265	2	84
Limestone										
Adipic acid-enhanced										
Case 1[g]	20	NA	11.3	--	753	68	-90	+233	5	130
Case 2[e]	5-10	NA	4.9	--	586	58	-100	+292	5	110
MgO-promoted[h]	10	20	8.4	--	1378	64	-90	+322	3	130
Daily										
Dual alkali										
FMC Corp.[e]	12	20	--	12.8	51	22	-49	+40	6	45
CEA/ADL[f]	5	10	--	7.0	91	31	-60	+86	6	62
Limestone										
Adipic acid-enhanced										
Case 1[g]	20	NA	--	10.4	28	57	-70	+165	22	120
Case 2[e]	5-10	NA	--	4.7	21	41	-74	+97	19	88
MgO-promoted[h]	10	20	--	8.4	49	56	-73	+156	16	110

[a]Equal to 100% minus the % control efficiency.

[b]Confidence intervals extending beyond -100% should be truncated at -100%.

[c]Based on Eq. B.4 (see Sec. B.2.1).

[d]Based on Eq. B.5 (see Sec. B.2.2).

[e]Mean values given are for two units or two experimental design conditions.

[f]Combustion Equipment Associates/Arthur D. Little, Inc.

[g]Mean values are for four experimental design conditions.

[h]American Air Filters, Inc., system. MgO = magnesium oxide.

NA = not available.

mean values are rather small, at least for the hourly averaging period, reflecting the effect of the large number of samples taken.

The relative 95% confidence intervals for individual averaging periods range from a minimum interval of about ±60% to a maximum interval of -100% to +130% for the hourly averaging period, and from a minimum interval of ±40% to a maximum interval of -100% to ±120% for the daily averaging periods.

Table 3.4 also shows the manufacturer's design and guaranteed penetration factors for comparison with the means of the measured penetration factors. The measured factors differ from the design factors by about 1-9% and from the guaranteed factors by about 3-12%, suggesting that the potential level of uncertainty due to estimation errors could also be substantial.

3.5 EMISSION FACTORS

The emission factors preferred for use in developing the NAPAP emissions inventories are as follows (listed in order of preference):[7]

1. Individual-source emission factors based on stack test results, other emission measurements, or material balances,

2. EPA's SCC-group mean emission factors (AP-42 emission factors),[11]

3. Emission factors other than the AP-42 emission factors, and

4. Default emission factors to be developed by EPA.[7]

Of these possibilities, the AP-42 emission factors are expected to be used for most of the emission calculations. Thus, our efforts concentrated on compiling and developing uncertainty values for these factors. A limited amount of background data on individual-source emission factors was also analyzed to gain some idea of the uncertainties associated with these factors.

For compilation of existing uncertainty values, the emission source data (i.e., emission factors for individual sources) that were used and reported by PEDCo[3] were evaluated. For development of new uncertainty values, the emission source data used by EPA in developing its latest AP-42 emission factors[5] and other emission source data readily available in the open literature were analyzed. The results of these data analyses are presented below for major combustion point sources.

3.5.1 SO_x Emission Factors

Table 3.5 lists the mean emission factors, number of data points on which each mean factor was based, variability in terms of CV, and relative extreme values for several sets of SO_x emission source data. The AP-42 emission factors are also listed for

TABLE 3.5 Variability of Combustion Point Source Emission Factors for SO_x

Fuel	Data Set	SCC Boiler Type[a]	Emission Factors[b] (lb/ton) AP-42	Emission Factors[b] (lb/ton) Mean[c]	No. of Samples	CV (%)	Relative Extreme Values (%) Low	Relative Extreme Values (%) High	Data Source (Ref.)	Data Form
Coal	1	Electric utility, bituminous coal	39S							
		Wet bottom		38.3S	7	25	-36	+49	3	D.3
		Dry bottom		38.3S	8	23	-36	+49	3	D.3
		Cyclone		36.7S	4	9	-13	+5	3	D.3
	2	All boiler types		33.3S	152	18	-61	+23	12	D.4
	3	All pulverized coal; spreader, and overfeed stokers								
		Bituminous coal	39S	39.3S	49	20	-49	+50	13	D.5
		Subbituminous coal	35S	35.0S	15	25	-38	+46	13	D.5
		underfeed stoker: bituminous coal	31S	31.0S	9	28	-38	+53	13	D.6
	4	Electric utility[d]	39S							
		Vertical		43.5S	5	15	-15	+24	14	D.7
		Front wall		56.0S	5	23	-37	+25	14	D.7
		Corner		43.9S	5	5	-7	+5	15	D.7
		Horizontally opposed		39.9S	4	5	-6	+7	15	D.7
Residual oil	5	All boiler types: utility, industrial, and commercial-institutional	157S[e]	138S	97	23	-86	+17	16	D.8

[a] See App. D data forms for SCC numbers.

[b] S refers to the sulfur content in weight %.

[c] Emission factors greater than 40.0S represent fuel sulfur conversion to sulfur oxides over 100%, indicating the presence of errors associated with emission measurements and/or sulfur content data used in deriving the emission factors.

[d] All values reported are based on individual measurements at single sources.

[e] In lb/10^3 gal.

comparison. The SO_x emission factors, e.g., 39S lb of SO_2 per ton of bituminous coal burned, consist of two parts: the numerical part, i.e., 39 in this case, and the sulfur content part, S (in weight %). The numerical part is the SCC-group mean fractional conversion factor for converting the fuel sulfur to SO_x (per unit weight % of the fuel sulfur content), and S is the individual-source mean sulfur content for the period for which emissions are calculated, e.g., an annual or hourly period. Therefore, the uncertainty due to variability in the fractional conversion factor and that due to the variability in fuel sulfur content must be accounted for separately. Since uncertainty in the mean fuel sulfur content has already been discussed in Sec. 3.2, only the uncertainty in the fractional conversion factor is discussed here. Variability in the fractional

conversion factor results from the differences in boiler type and size and in the operating and ambient conditions associated with the source emissions test data from which the emission factors are derived. Because fuel sulfur content data are used in deriving the SO_x emission factor from source emissions test data, uncertainty due to the variability in fuel sulfur content could be embedded in the numerical part of the SO_x emission factor. Such uncertainty could be rather large.

The first set of data in Table 3.5 is based on the data used by PEDCo[3] to develop SCC-group mean SO_x emission factor uncertainty values for several different types of coal-fired electric utility boilers. Examination revealed that the data set used to develop the uncertainty values for a given SCC category included data for other SCC categories and for multiple-SCC categories. When these data were eliminated, only a couple of individual-source emission data points remained for some boiler types. Thus, this first data set and, consequently, any uncertainty data developed from it are of poor quality as far as any individual SCC category is concerned.

The second set of data[12] in Table 3.5 covers all types of coal combustion sources, ranging from domestic stoves to large electric utility boilers with the exception of locomotives. The mean emission factor of 33.3S for this data set is considerably lower than the AP-42 emission factors for all other boiler categories listed in the table except underfeed stoker boilers. The mean was based on a large number of source data points (152) and has a relatively small data spread. However, the mean covers a broad range of sources and is therefore less specific than the AP-42 emission factors.

The third data set is based on data used in developing AP-42 emission factors for multiple-SCC categories.[13] The number of data points used in developing the mean emission factor for the pulverized coal-fired boilers and spreader and overfeed stoker boilers burning bituminous coal is relatively large (49), and the CV value is 20%. Fewer data points were used to develop the mean emission factors for the same category of boilers, but burning subbituminous coal, and for underfeed stoker boilers (15 and 9 data points, respectively), and the CV values are somewhat greater (25% and 28%, respectively).

The fourth data set is based on individual-source emission factors for coal-fired electric utility boilers. The numbers of data points are quite limited and the mean emission factor for each source is substantially different from the multiple-SCC-group mean AP-42 emission factors. The source emission data for several of the electric utility boilers[14,15] were obtained at two specific load levels (75% and 100%). Thus, the data set does not adequately represent the distribution of the load or other operating and ambient conditions. Even under these limited conditions, the variability in the data sets for certain boiler types is comparable to the variabilities for multiple-SCC-group mean AP-42 emission factors.

The fifth data set in Table 3.5 was drawn from the same data source[16] as that used for the PEDCo uncertainty data.[3] The number of data points is relatively large (97), and the CV value is 23%. However, the mean emission factor for this data set is different from the AP-42 emission factor, indicating that the latter has been revised. Thus, the variability values derived from this data set may not be appropriate for the revised AP-42 emission factor (as far as application to the 1985 NAPAP emissions

inventory is concerned). However, engineering judgment was used in developing the new AP-42 emission factors for residual oil-fired boilers, and therefore an uncertainty value based on data variability cannot be derived.

3.5.2 NO_x Emission Factors

Table 3.6 contains data on NO_x emission factors similar to those provided in Table 3.5 for SO_x. The data used by PEDCo[3] in developing uncertainty values for SCC-group mean NO_x emission factors for electric utility boilers also included data for other SCCs and for multiple-SCC categories. Elimination of these data again left only a few data points. Therefore, the remaining PEDCo data points were supplemented by more-recent data compiled by KVB, Inc.,[17] for use in developing uncertainty values for SCC-group mean NO_x emission factors for electric utility and industrial boilers.

The first two data sets in Table 3.6 are based on these composite data. The mean emission factors computed are quite comparable to the corresponding AP-42 emission factors except in a few cases. The average CV values for the electric utility and industrial boilers are similar to each other (27% and 23%, respectively), but they may not be reliable measures of variability for the source population due to the small amounts of source emission data used.

The third data set in Table 3.6 is based on the source emission data used in developing the AP-42 emission factors.[13] Except for a couple of cases, the amount of source emission data is substantially greater than for either of the first two data sets. One reason for this abundance of data is that the AP-42 emission factors are developed for multiple-SCC categories, including utility, industrial, and commercial-institutional boilers. The CV values range from about 20% to 40% with only one exception (over 70%). The spread of relative extreme values is somewhat greater than for the first two sets of data.

The fourth data set in Table 3.6 is based on individual-source emission factors for coal-fired electric utility boilers.[14,15] The amount of source emission data is quite limited, and the mean emission factors are substantially different from the multiple-SCC AP-42 emission factors. The CV values and relative extreme values have ranges similar to those for the AP-42 emission factors.

The last two data sets in Table 3.6 are for SCC-group mean emission factors for residual oil- and natural gas-fired boilers. The amount of source emission data is not very large, except for the residual oil-fired electric utility boilers. The CV values for the mean NO_x emission factors for residual oil-fired boilers are comparable to those for coal-fired boilers, while the CV values for natural gas-fired boilers are greater. The ranges of relative extreme values for these two data sets are in general slightly greater than those for coal-fired boilers.

The mean emission factors for these last two data sets are quite different from the AP-42 emission factors. Variability data need to be developed, based on the data used in developing the current AP-42 emission factors, for use in the NAPAP emissions inventory uncertainty assessment.

TABLE 3.6 Variability of Combustion Point Source Emission Factors for NO_x

Fuel	Data Set	SCC Boiler Type[a]	Emission Factors[b] (lb/ton) AP-42	Emission Factors[b] (lb/ton) Mean[c]	No. of Samples	CV (%)	Relative Extreme Values (%) Low	Relative Extreme Values (%) High	Data Source (Ref.)	Data Form
Coal	1	Electric utility, bit. coal								
		Wet bottom	34	34.8	5	35	-54	+37	3, 17	D.9
		Dry bottom	21	20.1	16	32	-45	+56	3, 17	D.10
		Cyclone	37	39.3	7	30	-25	+38	3, 17	D.11
		Tangential	15	14.5	17	10	-19	+19	17	D.12
	2	Industrial, bit. coal								
		Wet bottom	34	22.7	2	38	-27	+27	17	D.13
		Dry bottom	21	14.3	3	29	-21	+33	17	D.13
		Cyclone	37	26.0	1	--	--	--	17	D.13
		Tangential	15	14.5	3	14	-15	+11	17	D.13
		Spreader stoker	14	15.2	10	18	-31	+25	17	D.14
		Overfeed stoker	7.5	6.5	3	14	-11	+16	17	D.14
		Underfeed stoker	9.5	9.6	4	13	-10	+17	17	D.14
	3	All single wall, opposed and vertical; bit. and subbit. coal								
		Utility, industrial, and CI								
		Wet bottom	34	33.8	2	71	-51	+49	13	D.15
		Dry bottom	21	21.1	28	39	-55	+111	13	D.15
		Cyclone	37	36.5	7	24	-25	+49	13	D.16
		Tangential	15	14.7	29	29	-29	+57	13	D.17
		Spreader stoker	14	13.7	35	20	-49	+28	13	D.18
		Overfeed stoker	7.5	7.4	15	21	-32	+28	13	D.19
		Industrial and CI: underfeed stoker	9.5	9.5	8	23	-26	+42	16	D.20
	4	Electric utility								
		Vertical[c]	21	10.8	5	39	-52	+55	14	D.21
		Front wall[c]	21	23.4	5	19	-26	+24	14	D.21
		Corner[c]	21	15.2	4	15	-21	+14	15	D.21
		Horizontally opposed[c]	34	15.4	4	15	-14	+22	15	D.21
Residual oil	5	Electric utility								
		Normal (horizontal)	67[d]	105[d]	189	21	-48	+53	16	D.22
		Tangential	42[d]	50[d]	75	24	-58	+108	16	D.22
		Industrial and CI	67[d]	57[d]	17	78	-86	+135	16	D.23
Natural gas	6	Electric utility								
		General	550[e]	576[e]	13	68	-83	+136	3	D.24
		Tangential	275[e]	290[e]	14	35	-62	+53	3	D.24
		Industrial	140[e]	296[e]	4	74	-52	+109	3	D.25
		CI and domestic	100[e]	90[e]	14	103	-59	+292	3	D.26

[a] See App. D data forms for SCC numbers. Abbreviations in this column are as follows: bit. = bituminous, subbit. = subbituminous, and CI = commercial-institutional.

[b] In lb/ton except where otherwise noted. Data expressed in terms of $lb/10^6$ Btu were converted to lb/ton based on the fuel heating value or on 12,000 Btu/lb for bituminous coal when the heating value was not available.

[c] All values reported are based on individual measurements at single sources.

[d] In $lb/10^3$ gal.

[e] In $lb/10^6$ ft^3.

3.5.3 Confidence Intervals about the Mean Emission Factors

The uncertainties in using an SCC-group mean SO_x or NO_x AP-42 emission factor to represent either the true mean emission factor for the entire source population (i.e., for one or multiple SCC categories) or the mean emission factor for k sources of the same SCC were calculated assuming normal distribution and are listed in Table 3.7. The degree of uncertainty is expressed in both cases in terms of relative 95% confidence intervals. The other data in Table 3.7 are summarized from Tables 3.5 and 3.6.

The relative 95% confidence interval for the true mean emission factor, i.e., population mean, ranges from about ±5% to about ±20% for both the AP-42 SO_x and NO_x emission factors, except for one case involving two data points (-100% to +640%). When the AP-42 emission factor is used to represent a single source, most of the relative 95% confidence intervals range from about ±40 to ±80%. One exception, which is based on only two data points, is the relative 95% confidence interval for the NO_x emission factor for coal-fired wet-bottom boilers (-100% to +1100%). As the number of sources increases in the group that the AP-42 emission factor is used to represent, the span of the relative 95% confidence interval decreases, at first rapidly, then tapering off after about 10 sources.

The uncertainty associated with using an individual-source mean emission factor to represent the true emission factor for that source under a particular combination of conditions can be assessed using the individual-source emissions test data. The operating conditions of an electric utility boiler may change quite rapidly or remain steady for a long time, depending on the type of boiler and the time of day. If one assumes that these operating conditions, and consequently the emission factor, remain steady for a given period of time, e.g., 1 hr, then one can calculate the 95% confidence interval for that hour's emission factor (assuming normal distribution) as a rough measure of the uncertainty in using that emission factor to represent the emission factors under different operating conditions. The resulting confidence intervals, using the individual-source emissions data presented in Tables 3.5 and 3.6, are given in Table 3.8 for SO_x and NO_x emission factors for various types of coal-fired electric utility boilers. The relative 95% confidence intervals for the emission factor for a testing period, e.g., 1 hr, range from a minimum interval of ±15% to a maximum interval of ±70% for SO_x emission factors and from a minimum interval of ±50% to a maximum interval of -100% to +120% for NO_x emission factors. These confidence intervals are of the same order of magnitude as those for the AP-42 emission factors when they represent a single source.

3.6 TEMPORAL ALLOCATION FACTORS

Temporal allocation factors for the 1980 NAPAP emissions inventory were developed by GCA.[18] For electric utility point sources (i.e., power plant generating units), the seasonal temporal allocation factors were developed based on power generation statistics obtained from the 1979 *Energy Data Reports* published by the U.S. Department of Energy, and the hourly temporal allocation factors were based on hourly fuel use data previously collected in the Electric Power Research Institute's Sulfate Regional Experiment (SURE) program.[19] For other point sources, these factors were

TABLE 3.7 Variability and Uncertainty Values for AP-42 Emission Factors for SO_x and NO_x

Pollutant, SCC Boiler Type[a]	Emission Factor (lb/ton)	No. of Samples	CV (%)	Relative 95% Confidence Interval (±%)[c]: For the Parent Population Mean[d]	For the Mean for One or More Sources[e]: 1	5	10	50	100	500
SO_x										
All pulverized coal, spreader, and overfeed stokers										
Bituminous coal	39S[b]	49	20	6	41	19	14	8	7	6
Subbituminous coal	35S[b]	15	25	14	55	28	22	16	15	14
Underfeed stoker: bituminous coal	31S[b]	9	28	22	68	36	30	23	22	22
NO_x										
All single wall, opposed and vertical; bituminous and subbituminous coal										
Utility, industrial, and CI										
Wet bottom	34	2	71	640	1,100	750	700	650	640	640
Dry bottom	21	28	39	15	81	39	29	19	17	16
Cyclone	37	7	24	22	63	34	29	29	23	22
Tangential	15	29	29	11	60	29	22	14	13	11
Spreader stoker	14	35	20	7	41	19	15	9	8	7
Overfeed stoker	7.5	15	21	12	47	23	18	13	12	12
Industrial and CI: underfeed stoker	9.5	8	23	19	58	31	26	21	20	19

[a] In this column, CI = commercial-institutional.

[b] S refers to the sulfur content in weight % based on measurements.

[c] Confidence intervals that extend beyond -100% should be truncated at -100%.

[d] Based on Eq. B.4 (see Sec. B.2.1).

[e] Based on Eq. B.5 (see Sec. B.2.2).

based on operating rate information contained in annual inventory records. For area sources, these factors were based on published activity statistics (e.g., gasoline sales).

Uncertainty values based on measured data have not been developed to date for the temporal allocation factors. However, uncertainty values could be estimated by analyzing the data that GCA used to develop these factors.

To roughly ascertain the magnitude of uncertainty values for temporal allocation factors for the electric utility sector, the 1985 hourly net generation data for each unit in the 13 TVA generating stations (excluding hydroelectric units) were analyzed.[6] The

TABLE 3.8 Variability and Uncertainty Values for Individual-Source Mean Emission Factors for SO_x and NO_x

Pollutant, Type of Utility Boiler	Emission Factor (lb/ton)	No. of Samples	CV (%)	Relative 95% Confidence Interval (±%)[a] For the Parent Population Mean[b]	Relative 95% Confidence Interval (±%)[a] For the Mean for One Sampling Period[c,d]
SO_x					
Vertical	43.5S[e,f]	5	15	19	46
Front wall	56.0S[e,f]	5	23	28	70
Corner	43.9S[e,f]	5	5	6	15
Horizontally opposed	39.9S[e]	4	5	8	18
NO_x					
Vertical	10.8	5	39	48	120
Front wall	23.4	5	19	24	58
Corner	15.2	4	15	24	53
Horizontally opposed	15.4	4	15	25	53

[a]Confidence intervals extending beyond -100% should be truncated at -100%.

[b]Based on Eq. B.4 (see Sec. B.2.1).

[c]Based on Eq. B.5 (see Sec. B.2.2).

[d]For example, 1 hr.

[e]S refers to the sulfur content in weight % based on measurements.

[f]Any SO_x emission factors greater than 40.0S imply that the conversion of sulfur to sulfur oxides is > 100%.

amount of electricity generated is a surrogate for fuel throughput. Use of a surrogate value would introduce additional uncertainty. However, such additional uncertainty was not investigated in this interim report. Of the 64 units in total, 59 are coal-fired and 5 are nuclear. Data on the nuclear units were included on the assumption that their generation patterns would be similar to those of large coal-fired baseload units. Small negative hourly net generation values were treated as zero in the analysis.

Uncertainty values were estimated for the electric utility sector's quarterly and mean hourly allocation factors, assuming normal distribution of the individual-unit

factors. As defined in this report, a *quarterly allocation factor* is used for disaggregating annual emissions to emissions for a specific quarter (i.e., first, second, third, or fourth). A *mean hourly allocation factor* is used for disaggregating the resulting specific-quarter emissions to a mean value for a specific hour in that quarter (e.g., for all of the 1-2 p.m. periods in the first quarter). The derivations are further explained below.

The sector's mean quarterly allocation factors (Table 3.9) were obtained by averaging quarterly allocation factors derived for each of 63 units. These latter factors were derived by dividing each unit's net generation during each quarter by the unit's annual net generation (resulting in four quarterly allocation factors for each unit). Table 3.9 gives the variability in these individual-unit quarterly allocation factors in terms of CV values, which range from about 30% to 65%. Also provided are the uncertainty estimates associated with the mean quarterly allocation factors for the sector. The relative 95% confidence interval for the population mean quarterly allocation factors ranges from ±8% to ±17%. The relative 95% confidence interval when the sample mean is used to represent k sources ranges from a minimum interval of ±60% to a maximum interval of -100% to +130% when k = 1, decreases to ±20% to ±45% when k = 10, and then tapers off as k increases further.

The mean hourly allocation factors for the electric utility sector were obtained by averaging the mean hourly allocation factors for each unit. These latter factors were

TABLE 3.9 Variability and Uncertainty Values for Mean Quarterly Allocation Factors in the Electric Utility Sector[a]

Quarter	Mean Quarterly Allocation Factor	Number of Units	CV (%)	Relative Extreme Values (%)		Relative 95% Confidence Interval (±%)[c]					
				Low[b]	High	For the Parent Population Mean[d]	For the Mean for One or More Units[e]				
							1	5	10	50	100	500
1	0.254	63	65.7	-100	+294	17	130	61	45	25	21	18
2	0.275	63	39.3	-100	+83	10	79	36	27	15	13	11
3	0.238	63	30.8	-100	+58	8	62	29	21	12	10	8
4	0.233	63	44.5	-100	+148	11	90	41	30	17	14	12

[a]Based on the 1985 hourly net electricity generation data for 64 TVA units. Only 63 units contributed data in each quarter.

[b]The values listed result from the fact that small negative net generation values were treated as zero.

[c]Confidence intervals extending beyond -100% should be truncated at -100%.

[d]Based on Eq. B.4 (see Sec. B.2.1).

[e]Based on Eq. B.5 (see Sec. B.2.2).

derived for each quarter, category of day (i.e., weekday, Saturday, or Sunday), and hour of the day. The method was to divide each unit's net generation for each specific hour (e.g., 1-2 p.m. on April 25, 1985) by the unit's net generation for the corresponding quarter. The resulting specific-quarter-to-specific-hour allocation factors were then averaged by hour of the day for each quarter and category of day. The results are summarized in Table 3.10. These data are provided only for the hours of the day with the average, greatest and least variability in individual-unit hourly allocation factors. Also provided are the uncertainty estimates associated with the electric utility sector mean hourly allocation factors, along with the variability among unit hourly allocation factors in terms of CV.

In general, there seems to be no significant variation in these hourly values among calendar quarters and day-of-the-week categories. The CV values in Table 3.10 range from 5% to 187%, with the mean of these values ranging from 13% to 22%, depending on the category of day. The relative 95% confidence interval for the population mean hourly allocation factor ranges from ±3% to ±6% for the hours with average variability but has a maximum range of ±1.2% to ±48%. The relative 95% confidence interval for the mean hourly allocation factor for k sources range from ±25% to ±44% for the hours with average variability when k = 1. The smallest interval is ±9% and the largest is -100% to +380%.

It is possible to use the *mean value* for a unit's net generation for a specific hour in a given category of day and a given quarter to represent the unit's *actual* net generation for the same hour on a specific date. In such applications, the mean value is subject to uncertainty due to the unit's day-to-day variability in each hour's net generation. To estimate this uncertainty for the electric utility sector, the mean values and variability (in terms of CV) in the unit hourly net generation data were determined and analyzed for each quarter, day-of-the-week category, and hour of the day. Table 3.11 lists the data analyzed and the resulting uncertainty estimates (derived assuming normal distribution of the unit hourly net generation data). This information is provided for the units and hours of the day with the average, greatest, and least variability in individual-unit hourly net generation.

In general, no significant variations in these values are apparent among different calendar quarters and day-of-the-week categories. The CV values in Table 3.11 range from less than 2% to over 800%, with the mean ranging from 77% to 92%, depending on the category of day. The relative 95% confidence interval for the population mean hourly net generation ranges from ±19% to ±56% for the units and hours with average variability, but has a maximum range of less than ±1% for the smallest interval to -100% to +220% for the largest interval. The relative 95% confidence intervals for the mean of k hours ranges from a minimum interval of -100% to +160% to a maximum interval of -100% to +210% for the units and hours with average variability when k = 1. Extreme values range from a minimum interval of ±3% to a maximum interval of -100% to +1600%.

TABLE 3.10 Selected Variability and Uncertainty Values for Mean Hourly Allocation Factors in the Electric Utility Sector[a]

Variability Case	Type of Hour: Category of Day	Type of Hour: Quarter	Type of Hour: Hour of the Day	Mean Hourly Allocation Factor	No. of Samples	CV (%)	Relative 95% Confidence Interval (±%)[b]: For the Parent Population Mean[c]	For the Mean for One or More Units[d]: 1	5	10	50	100	500
Hours with average variability	Weekday	--[e]	--[e]	--[e]	60	12.6	3.3	25	12	9	5	4	3
	Saturday	--[e]	--[e]	--[e]	60	17.1	4.4	35	16	12	7	6	5
	Sunday	--[e]	--[e]	--[e]	60	21.6	5.6	44	20	15	8	7	6
Hours with the greatest variability	Weekday	4	11	0.00066	60	187	48	380	170	130	72	61	48
	Saturday	4	3	0.00041	60	23.8	6.1	48	22	16	9	8	7
	Sunday	4	16	0.00033	60	34.8	9.0	70	32	24	13	11	10
Hours with the least variability	Weekday	3	22	0.00048	60	4.5	1.2	9	4	3	2	1	1
	Saturday	3	22	0.00044	60	11.3	2.9	23	11	8	4	4	3
	Sunday	3	12	0.00046	60	11.9	3.1	24	11	8	5	4	3

[a]Based on the 1985 hourly net generation data for 64 TVA units.

[b]Confidence intervals extending beyond -100% should be truncated at -100%.

[c]Based on Eq. B.4 (see Sec. B.2.1).

[d]Based on Eq. B.5 (see Sec. B.2.2).

[e]Because the average variability, i.e., CV, is an artificially calculated mean value, there is no corresponding hour of the day or quarter, nor is there a mean hourly allocation factor.

TABLE 3.11 Selected Variability and Uncertainty Values for Specific-Hourly Net Generation Data for Individual Units in the Electric Utility Sector[a]

Variability Case	Type of Hour: Category of Day	Type of Hour: Quarter	Type of Hour: Hour of the Day	No. of Samples[b]	CV (%)	Relative 95% Confidence Interval (±%)[c]: For the Parent Population Mean[d]	Relative 95% Confidence Interval (±%)[c]: For the Mean for 1 Hour[e]
Units and hours with average variability	Weekday	--[f]	--[f]	65	77	19	160
	Saturday	--[f]	--[f]	13	87	53	200
	Sunday	--[f]	--[f]	13	92	56	210
Units and hours with the greatest variability	Weekday	4	10, 11	66	812	200	1,600
	Saturday	1,2,4	1-24[g]	13	361	220	820
	Sunday	1,2,4	1-24[g]	13	361	220	820
Units and hours with the least variability	Weekday	1	21	64	1.6	0.4	3
	Saturday	2	17	13	1.4	0.9	3
	Sunday	1	12	13	1.3	0.8	3

[a]Based on the 1985 hourly net generation data for 64 TVA units.

[b]That is, the number of hourly periods in the quarter (e.g., number of weekday 1-2 p.m. periods) on which the uncertainty calculations were based.

[c]Confidence intervals extending beyond -100% should be truncated at -100%.

[d]Based on Eq. B.4 (see Sec. B.2.1).

[e]Based on Eq. B.5 (see Sec. B.2.2).

[f]Because the average variability is an artificially calculated mean value, there is no corresponding quarter or hour of the day.

[g]Except for hours 1-5 during the second quarter, and hours 1 and 2 during the fourth quarter.

4 ILLUSTRATIVE EMISSIONS UNCERTAINTY ESTIMATES

To illustrate the usefulness of the findings summarized in earlier sections, estimates were made of the uncertainties in the SO_x and NO_x emissions estimates for coal-fired electric utility boiler units. Two cases were used: a single unit and a group of 100 identical units. Besides annual emissions, three levels of temporal disaggregation of these emissions estimates were examined. The mathematical procedures and all necessary assumptions are presented in App. C, and the results are summarized in Table 4.1. These estimates take into account the uncertainty associated with each of the following EEPs:

1. Annual activity level,
2. SO_x and NO_x emission factors,
3. Coal sulfur content,
4. Two temporal allocation factors:
   - Quarterly (for disaggregating annual emissions to values for each quarter) and
   - Mean hourly (for disaggregating specific-quarter emissions to a mean value for a specific hour, e.g., 1-2 p.m., in the quarter), and
5. Hourly activity level correction factor* (for using a mean hourly emissions value to represent the emissions for the same hour on a specific date, e.g., 1-2 p.m. on April 25, 1985).

Therefore, these uncertainty estimates apply to annual and temporally allocated emissions estimated by procedures similar to those used for the detailed 1980 and 1985 NAPAP emissions inventories. They do not apply to emissions estimates such as those made during a field test where real-time emissions (or emission factors, coal sulfur contents, and activity levels) may be measured for individual sources. (Even when individual-source EEPs are measured directly, there are many remaining causes of variability, which may not be detected unless the measurements are made continuously. These causes of variability include changes in load, operating conditions, and fuel characteristics.)

Listed specifically in Table 4.1 are the relative 95% confidence intervals for the annual and temporally allocated emissions. Two sets of these intervals have been derived

*The expected value of this factor is 1, with the uncertainty involved in using a mean hourly emissions value to represent the emissions for the same hour on a specific date.

TABLE 4.1 Uncertainty Estimates for SO_x and NO_x Emissions from One or a Group of 100 Identical Coal-Fired Electric Utility Boiler Units[a]

Pollutant, Period (col. 1)	Relative 95% Confidence Interval[b] (±%)		
		Emissions from 100 Units	
	Single-Unit Emissions (col. 2)	Product-Derived Approach (col. 3)	Aggregate-Derived Approach[c] (col. 4)
SO_x			
Year	41	7	4
Quarter	101[d]	17	10
Hour			
Mean	106[d]	17	11
Individual	228[d]	25	23
NO_x			
Year	70	15	7
Quarter	118[d]	21	12
Hour			
Mean	123[d]	22	12
Individual	234[d]	28	23

[a] No emission control is assumed.

[b] Expressed as ±% of the emissions estimate for each period.

[c] This approach neglects the uncertainty due to an insufficient amount of measurement data used in developing mean EEP values.

[d] Confidence intervals that extend beyond -100% should be truncated at -100%.

for the 100-unit case, one based on the product-derived approach (column 3), and the other on the aggregate-derived approach (column 4). These approaches were described in Sec. 1. Briefly for this illustration, in the product-derived approach, the total emissions for 100 units are estimated by multiplying the relevant EEPs for the 100 units and, in the aggregate-derived approach, emission estimates made for individual units are summed to compute the total emissions.

As an example of the results shown in Table 4.1, SO_x emissions estimates that are based on the mean specific-hourly operation of 100 identical units are subject to a ±17% error, with 95% confidence (column 3, product-derived approach). The results for this case are symmetrical about the mean emissions estimates. For the single-unit estimates, the results are not symmetrical except for the emissions computed on an annual basis. The uncertainty associated with NO_x emissions estimates (expressed as relative 95% confidence intervals), based on the operation of a single unit during one specific hour, ranges from -100% to about +230% of the emissions estimates (column 2). For the group emissions, the aggregate-derived approach yields smaller uncertainty intervals than the product-derived approach in all cases. The reason for the difference is that the aggregate-derived approach assumes that the values for SCC-group mean EEPs have been estimated for each individual unit, and therefore fails to treat the uncertainty due to an insufficient amount of measurement data used in developing mean EEP values, which leads to variability in knowing the true population mean.

Since the uncertainty estimates listed in Table 4.1 are based on a number of assumptions and an approximate methodology, they should not be considered as accurate estimates, but as rough approximations for illustrative purposes only. Key assumptions are (1) that variables in the emissions equation are independent and normally distributed (which is not true in some cases, especially specific-hour variation for a single unit) and (2) that the product of normally distributed random variables is also normally distributed. Although the latter assumption is reasonable when only two variables are involved, procedures must be developed for dealing with cases involving many variables and combinations of distribution types (e.g., normal and lognormal distributions) in the same equation.

5 SUMMARY OF RESULTS

In this report, the sources of uncertainties associated with EEPs were analyzed, uncertainty values for selected EEPs were identified through a literature review, and additional uncertainty values were developed using available basic data.

Three basic categories of EEPs were identified:

1. *Individual-source (or -county) values*: the annual activity level and spatial disaggregation factor,

2. *Individual-source mean values*: the mean sulfur or ash content of fuel, mean hourly activity level, and mean emission control equipment penetration factor,

3. *Mean values for SCC groups or individual counties*:

   - For single- or multiple-SCC groups: the AP-42 mean emission factor, mean species speciation factor, and mean temporal allocation factor, and

   - For individual counties: the mean sulfur or ash content of fuel for area sources.

A variety of errors result when the values for these EEPs are based on a limited number of samples and are used to represent (1) the parent population means, (2) other averaging periods, or (3) one or more sources. Errors in these cases are called representation errors. Other sources of errors in EEPs include measurement errors, data processing errors (which include derivation errors, rounding errors, and data transfer errors), and estimation errors. Definitions of these errors and their generation processes have been delineated in this interim report.

In compiling existing uncertainty values for EEPs, only those based on measured, reported, or survey data were considered, and those based on engineering judgment or analysis were excluded. The EEPs for which uncertainty values have been identified or developed in this study were limited to the following cases:

1. Annual activity level for point sources,

2. Coal sulfur content for point sources,

3. FGD system penetration factors for the electric utility sector,

4. SO_x and NO_x emission factors for fuel combustion point sources, and

5. Temporal allocation factors for the electric utility sector.

For the annual activity level, only rounding errors with embedded measurement errors have been reported as the measure of uncertainty. For the remaining four parameters, which are either individual-source mean values or SCC-group mean values, this report provides the mean value and associated basic data, including the number of data points on which the mean is based, variability or spread of the data points in terms of CV, and relative extreme values. The basic data for these EEPs represent a variety of data distribution types. Some distributions appear to be relatively simple and close to the normal distribution, while others are not. To roughly ascertain the magnitude of the uncertainty associated with these parameters in actual applications, three measures of uncertainty associated with the mean values were computed, assuming normal distributions:

1. *Relative 95% confidence interval for the population mean*: for the first type of representation errors, i.e., those that occur when the mean of a sample population is used to represent the parent population mean,

2. *Relative 95% confidence interval around the individual source mean when it is used to represent one or a group of values*: for the second type of representation errors, i.e., those that occur when an individual-source mean is used to represent the mean for other averaging periods (e.g., any 1-hr period or any group of 1-hr periods),

3. *Relative 95% confidence interval around the SCC-group mean when it is used to represent one or a group of values*: for the third type of representation errors, i.e., those that occur when an SCC-group mean is used to represent the mean for any one or more sources or units within the same SCC category.

For the coal sulfur content at point sources, the magnitude of the measurement error has also been examined.

The uncertainty values thus compiled or developed (assuming normal distribution and expressed as relative 95% confidence intervals) are summarized in Table 5.1. Typical or mean uncertainty values are given in the table to simplify discussion, but the ranges of values developed in this report for different cases are also given in parentheses for reference. The ranges represent the minimum and maximum values for those cases. Since rounding errors are qualitatively similar to the second type of representation errors listed above, and since measurement errors are embedded in the rounding errors and representation errors, the degree of uncertainty for different EEPs can be compared in terms of the three types of representation errors. Of the uncertainties due to these types of representation errors, those due to the second type (columns 3-5 in Table 5.1) are the largest, followed by those due to the third type (columns 6-8 in Table 5.1), then those due to the first type (column 2 in Table 5.1). For some EEPs, the relative 95% confidence intervals extend beyond -100%, an anomaly caused by assuming normal distribution for data distributions having standard deviations that are greater than half the mean values. The lower bound in such cases should be truncated at -100%. In those cases, the reliability of the upper bound is also subject to question.

TABLE 5.1 Summary of Uncertainty Values Identified or Developed in This Report for Selected EEPs[a]

EEPs by Category (col. 1)	Uncertainty Due to Representation Errors: When the EEP Represents the Parent Population Mean (col. 2)	When the EEP Represents the Mean for One Averaging Period: Hour (col. 3)	Day (col. 4)	Quarter (col. 5)	When the EEP Represents the Mean for k Sources[b]: k = 1 (col. 6)	k = 10 (col. 7)	k = 100 (col. 8)	Uncertainty Due to Rounding Errors (col. 9)	Uncertainty Due to Measurement Errors (col. 10)
Individual-source value: annual activity level								<5[c] (0.1-50)	
Individual-source mean values									
Coal sulfur content	0.5[d,e] (0.3-0.8)	45[d] (20-80)	15[d] (10-20)	2[d] (1-4)					7.5[f] (5-10)
FGD system penetration factor[g]	3[e] (2-5)	100 (60-130)	85 (40-120)						
Emission factor[h]									
SO_x from coal combustion	15 (5-30)	40 (15-70)							
NO_x from coal combustion	30 (25-50)	70 (50-120)							
Hourly activity level[i,j]	30 (0.5-200)	170 (3-1600)							
SCC-group mean values									
Emission factor[k]									
SO_x from coal combustion	15 (5-20)				50 (40-70)	20 (15-30)	15 (10-25)		
NO_x from coal combustion[l]	15 (5-600)				60 (40-1100)	25 (15-400)	20 (10-300)		

TABLE 5.1 (Cont'd)

EEPs by Category (col. 1)	Uncertainty Due to Representation Errors: When the EEP Represents the Parent Population Mean (col. 2)	When the EEP Represents the Mean for One Averaging Period: Hour (col. 3)	Day (col. 4)	Quarter (col. 5)	When the EEP Represents the Mean for k Sources[b]: k = 1 (col. 6)	k = 10 (col. 7)	k = 100 (col. 8)	Uncertainty Due to Rounding Errors (col. 9)	Uncertainty Due to Measurement Errors (col. 10)
Temporal allocation factor[i]									
Quarterly	12 (10-15)				90 (60-130)	30 (20-45)	15 (10-20)		
Hourly[j]	4 (1-50)				30 (10-380)	10 (3-130)	5 (2-60)		

[a]Each value listed is the mean for the range of values given in parentheses below it. Confidence intervals that extend beyond -100% should be truncated at -100%.

[b]Randomly selected from the parent population.

[c]Typical value based on industry conventions and survey data by PEDCo.[3] The range covers various SCC categories.

[d]Based on model-generated data using a first-order autoregressive model. The ranges are for 500-MWe and 100-MWe coal-burning boilers at a rate of 0.4 ton of coal/hr/MWe.

[e]For hourly mean values.

[f]For each measurement, based on the allowed limits of reproducibility for duplicate measurements in the ASME power test codes.

[g]For the electric utility sector. Values are based on test results for four commercial-scale power plant FGD systems.

[h]Based on 4 to 5 measurements at each of four types of boilers. Estimates assume that each measurement is an hourly average.

[i]For the electric utility sector, based on the 1985 hourly net generation for 64 TVA units.

[j]Values based on the variability in hourly net generation for each unit, by quarter, day-of-the-week category, and hour of the day.

[k]For the EPA AP-42 emission factors for various types of coal-fired boilers.

[l]In each case, mean values were computed excluding one extreme value.

The mean uncertainty values due to the first type of representation errors (column 2 in Table 5.1) range from less than ±1% to ±30% (the actual values are as small as less than ±1% and as large as -100% to +600%). The greatest uncertainty is associated with the utility mean hourly activity level (±30%), followed by source-specific NO_x and SO_x emission factors for coal-burning point sources (±30% and ±15%, respectively), then SCC-specific NO_x and SO_x AP-42 emission factors for coal-burning point sources (±15%), utility temporal allocation factors (about ±10% for the mean quarterly factor and ±4% for the mean specific-hourly factor), and the utility FGD system penetration factor (±3% for an hourly averaging period). The coal sulfur content shows very little uncertainty for this kind of error (±0.5%).

The mean uncertainty values due to the second type of representation errors (columns 3-5 in Table 5.1) range from a minimum relative 95% confidence interval of ±40% to a maximum interval of -100% to +180% for an individual source mean when it is used to represent a single hourly averaging period (the actual uncertainty values are as small as ±3% and as large as -100% to +1600%). The greatest uncertainty of this kind is exhibited by the utility mean hourly activity level (-100% to +180%), followed by the utility FGD system penetration factor (±100%), the source-specific NO_x and SO_x emission factors for coal-burning point sources (±70% and ±40%, respectively), and the coal sulfur content (±45%). Rounding errors for the annual activity level, also classified as belonging to this type of error, show the smallest uncertainty (typically <5%).

The mean uncertainty values due to the third type of representation errors (columns 6-8 in Table 5.1) range from ±20% to ±90% in terms of the relative 95% confidence interval about a sector or SCC-group mean when it is used to represent a single source or unit (the uncertainty values are as small as ±10% and as large as -100% to +1100%). The greatest uncertainty of this kind is exhibited by the utility mean quarterly allocation factor (±90%), followed by the NO_x and SO_x AP-42 emission factors for coal-burning point sources (±60% and ±50%, respectively) and the utility mean hourly allocation factor (±30%).

To illustrate the usefulness of these uncertainty estimates for the EEPs, estimates were made in Sec. 4 of the uncertainties associated with SO_x and NO_x emissions data for a single coal-fired electric utility boiler unit and a group of 100 identical such units, assuming no emission control devices. The estimates were based on the typical or mean uncertainty values associated with the pertinent EEPs. Approximate mathematical procedures and several simplifying assumptions were used. The relative 95% confidence intervals thus estimated for SO_x and NO_x emissions over various time periods were listed in Table 4.1. For a single unit, the relative 95% confidence intervals for SO_x emissions are quite large, ranging from about ±40% to an interval of -100% to about +230%, depending on the averaging period. The relative 95% confidence intervals for NO_x emissions are roughly comparable. The emissions uncertainty estimates become smaller for a group of units. For a group of 100 units, the relative 95% confidence intervals for SO_x emissions (based on the product-derived approach) range from ±7% to ±25%, depending on the averaging period, and the intervals for NO_x emissions range from ±15% to ±28%. For the 100-unit case, the aggregate-derived approach yields smaller uncertainty intervals than the product-derived approach in all cases. The reason for the difference is that the aggregate-derived approach assumes that the values for SCC-

specific EEPs have been estimated for each individual unit, and therefore fails to treat the uncertainty due to an insufficient amount of measurement data used in developing mean EEP values, which leads to variability in knowing the true population mean.

These illustrative, preliminary uncertainty estimates are rough approximations and apply to typical or average situations only, because they are based on typical or mean values for the magnitude of uncertainty associated with the pertinent EEPs. Depending on the situation, e.g., whether the units under consideration are baseload or intermediate-load units, or whether the time period of concern falls on a weekday or weekend, the emissions uncertainty estimates could be larger or smaller.

While these estimates are derived for the electric utility sector, they might suggest the order of magnitude of the uncertainty associated with SO_x and NO_x emissions from other point source combustion processes. Combustion point sources currently account for about 80% of the SO_x emissions and 50% of the NO_x emissions in the United States.

REFERENCES

1. *Solid Fuels: Power Test Codes*, American Society of Mechanical Engineers, New York, PTC 3.2-1954 (reaffirmed 1979).

2. *Diesel and Burner Fuels: Power Test Codes*, American Society of Mechanical Engineers, New York, PTC 3.1-1958 (reaffirmed 1979).

3. *Source Inventory and Emission Factor Analysis, Vols. 1 and 2*, prepared by PEDCo-Environmental Specialists, Inc., U.S. Environmental Protection Agency Report EPA-450/3-75-082-a,b (Sept. 1974).

4. *Compilation of Air Pollution Emission Factors*, U.S. Environmental Protection Agency Report AP-42, 2nd Ed. (April 1973).

5. *Compilation of Air Pollution Emission Factors, Vol. 1: Stationary Point and Area Sources*; and *Vol. 2: Mobile Sources*, U.S. Environmental Protection Agency Report AP-42, 4th Ed. (Sept. 1985).

6. Massey, C.L., Tennessee Valley Authority, Environmental Coordination, Chattanooga, Tenn., personal communication (April 9, 1986).

7. *NEDS/NAPAP Emission Inventory Workshop*, booklet for participants, U.S. Environmental Protection Agency, Washington, D.C. (Oct. 1985).

8. Kilgroe, J.D., *Coal Cleaning*, in Handbook on Air Pollution Control, John Wiley and Sons, New York (1983).

9. Cheng, B., et al., *Variability and Correlation in Raw and Clean Coal: Measurement and Analysis*, U.S. Environmental Protection Agency Report EPA-600/7-83-055 (April 1984).

10. Peduto, E.F., Jr., et al., *Statistical Analysis of Continuous SO_2 Emissions Data from Steam Generators: Vol. I, A.B. Brown Unit No. 1; Vol. II, Springfield City Utilities, Southwest Unit No. 1; Vol. III, Louisville Gas and Electric Company, Cane Run Unit No. 6*; and *Vol. IV, Allegheny Power Systems, Pleasants Unit No. 1*, prepared by GCA Corp. for the U.S. Environmental Protection Agency, Contract 68-02-3168, Task 53 (July 1983).

11. *NEDS Source Classification Codes and Emission Factor Listing*, U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, N.C. (Oct. 1985).

12. Smith, W.S., and C.W. Gruber, *Atmospheric Emissions from Coal Combustion — An Inventory Guide*, U.S. Public Health Service Report 999-AP-24 (1966).

13. Lahre, T., U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, N.C., personal communication (April 24, 1986).

14. Cuffe, S.T., et al., *Air Pollution Emissions from Coal-Fired Power Plants: Report No. 1*, J. of Air Pollution Control Assn., *14*:353-362 (1964).

15. Gerstle, R.W., et al., *Air Pollution Emissions from Coal-Fired Power Plants, Report No. 2*, J. of Air Pollution Control Assn., *15*:59-64 (1965).

16. Smith, W.S., *Atmospheric Emissions from Fuel Oil Combustion: An Inventory Guide*, U.S. Public Health Service Report 999-AP-2 (1962).

17. *Environmental Control Implications of Generating Electric Power from Coal; 1977 Technology Status Report, Appendix D: Assessment of NO_x Control Technology for Coal Fired Utility Boilers*, prepared by KVB Inc., Argonne National Laboratory Report ANL/ECT-3 (Dec. 1977).

18. Sellars, F.M., et al., *National Acid Precipitation Assessment Program Emission Inventory Allocation Factors*, prepared by GCA Corp., U.S. Environmental Protection Agency Report GCA-TR-84-171-G (March 1985).

19. *Emission Inventory for the SURE Region*, prepared by GCA Corp., Electric Power Research Institute Report EPRI EA-1913 (April 1981).

APPENDIX A:

LITERATURE REVIEW OF EEP AND EMISSIONS INVENTORY UNCERTAINTY DATA

APPENDIX A:

LITERATURE REVIEW OF EEP AND EMISSIONS INVENTORY UNCERTAINTY DATA

A.1 INTRODUCTION

Available uncertainty data on EEPs were examined to determine their applicability and relevance to the NAPAP emissions inventories uncertainty assessment. For this purpose, the following documents were reviewed:

1. *Source Inventory and Emission Factor Analysis, Vols. 1 and 2,* prepared by PEDCo-Environmental Specialists, Inc., U.S. Environmental Protection Agency Report EPA-450/3-75-082-a,b (Sept. 1974),

2. *Emission Inventory for the SURE Region,* prepared by GCA Corp., Electric Power Research Institute Report EPRI EA-1913 (April 1981),

3. *Emissions, Costs, and Engineering Assessment,* Work Group 3B, United States - Canada Memorandum of Intent on Transboundary Air Pollution (June 1982), and

4. *1980 NAPAP Emissions Inventory Uncertainty Workshop,* in Estimation of Uncertainty for the 1980 NAPAP Emissions Inventory, U.S. Environmental Protection Agency Report EPA-600/7-86-055 (Dec. 1986).

For convenience, these documents are referred to as the PEDCo report, the GCA/SURE report, the Work Group 3B report, and the NAPAP Workshop report, respectively.

Each report is briefly summarized below, with comments on the following topics, where appropriate:

1. Measures used to express the degree of uncertainty in the EEP data,

2. Methods used to estimate uncertainty data for each EEP, and

3. Quality and quantity of the basic data used to derive each uncertainty estimate.

A.2 PEDCo REPORT

The PEDCo report describes the work conducted by that company on the uncertainty in the data for various EEPs in the following emission equation:

$$E = (EF)(A)(PF) \tag{A.1}$$

where:

E = emissions from a point or area source of a given pollutant,

EF = emission factor, containing a term for fuel quality (ash or sulfur content), if appropriate,

A = activity level, e.g., material or fuel throughput, and

PF = penetration factor (1 - fractional control device efficiency) for a given pollutant.

The report describes the methods and data used by PEDCo in estimating EEP uncertainty values. Uncertainty estimates for various EEPs were made for individual source categories to the eight-digit SCC level. Uncertainty estimates for emission factors were made for particulate matter (PM), SO_x, NO_x, hydrocarbons (HC), and carbon monoxide, and for each of the following five bases for estimation, when appropriate:

1. Source test data or other emission measurements,
2. Material balances based on engineering knowledge and experience,
3. AP-42 emission factors compiled by EPA,
4. State or local emission factors that differ from the AP-42 emission factors, and
5. Guesses.

The measure adopted by PEDCo to express the degree of uncertainty was termed a *precision value*. Measured and reported data on the EEPs available as of 1974 and a variety of assumptions were used in making the uncertainty estimates. For emission factors derived from more than two observations, "precision" was mathematically defined as

$$P = \frac{S}{\bar{X}} \frac{1}{\sqrt{n}} \tag{A.2}$$

where:

S = sample standard deviation,

$\bar{X}$ = sample arithmetic mean, and

n = number of sample observations.

If one uses the following definition for CV,

$$CV = \frac{S}{\bar{X}} \tag{A.3}$$

then

$$P = \frac{CV}{\sqrt{n}} \tag{A.4}$$

When the number of observations n was large, the sample standard deviation was computed by using the equation:

$$S^2 = \frac{1}{n - 1} \sum (X_i - \bar{X})^2 \tag{A.5}$$

However, when n was equal to or less than 12, S was derived from the range R:

$$R = X_{max} - X_{min} \tag{A.6}$$

For other EEPs, including the emission factors derived using other methods, the precision value was obtained by dividing one of the following by the reported or an assumed value for a given EEP:

1. Range of rounding error, in the case of reported values such as the annual activity level,

2. Allowed limit of reproducibility, in the case of fuel sulfur or ash content,

3. Assumed range of values or assumed values, and

4. Standard deviation, assumed or derived from the range or allowed limit of reproducibility.

In this context, the definition of P is the same as that for CV in the last case above. In the first three cases, however, it could be considered as a certain multiple of CV.

A.3 GCA/SURE REPORT

The GCA/SURE report presents uncertainty estimates for the 1977 Sulfate Regional Experiment (SURE) regional annual emissions inventory. These estimates were

made by combining estimates for emissions from major individual-source categories and subcategories. Estimates of uncertainty in the EEPs in Eq. A.1 were used to derive estimates of uncertainty in the emissions of PM, SO_2, SO_x, nitric oxide (NO), nitrogen dioxide (NO_2), and three reactivity classes of HC from each of the source categories. Uncertainty estimates for species speciation factors were also used in deriving estimates of uncertainty for NO_x and HC emissions.

The measure adopted by GCA to express the degree of uncertainty was *accuracy*, or *error*, which was defined as the potential deviation from the true value. To determine this measure for a particular EEP, GCA used one of the following techniques:

1. Used the precision values or CV values reported by PEDCo,

2. Developed an average precision factor for a source category based on reported values for several subcategories,

3. Estimated a value from the precision values and CV values for a similar process,

4. Used a precision value as a base value and estimated an error value, or

5. Estimated an error value based on engineering judgment.

In estimating the uncertainty associated with emissions from an individual source category, GCA used the conventional method of combining the EEP errors quadratically. However, GCA also used an unconventional method of summing component errors linearly to combine the uncertainties associated with emissions from different source categories.

A.4 WORK GROUP 3B REPORT

The Work Group 3B report describes the procedures for estimating the uncertainties associated with the 1980 annual SO_x and NO_x emission inventories of the United States and Canada. The resulting estimates are also reported. Work Group 3B developed uncertainty estimates for EEPs at the state level for the U.S. inventories, and at the plant and provincial (or source region) level for the Canadian inventories. The procedures used in estimating both the emissions and the uncertainty values were somewhat different in each country.

The measure adopted by the Canadian group to express the degree of uncertainty was *precision*, which was described as being not "true" precision, but rather the best estimate of precision that could be obtained. However, a rigorous mathematical definition was not provided. The Canadian group assigned precision values to EEPs and other identified sources of errors in estimating SO_2 or NO_x emissions by applying engineering analyses based on published studies, previous experience, or engineering estimates. Error estimates for SO_2 were assigned on a plant-by-plant basis for large

emitters, while those for the remaining sources were assigned on a source category basis and applied to the provincial level. For NO_x emissions, error estimates were assigned on a source category basis to the provincial level, except for the power generation category; in that case, error estimates were assigned on a point source basis to the major plants.

The Canadian group also considered an error factor for spatial disaggregation of emissions. An additional error was introduced to all area source categories when provincial emission totals were prorated to emission source regions. Uncertainty estimates for NO_x species speciation were also considered by the Canadian group. These error estimates were included in the uncertainty values estimated for the individual sources and the source category emissions, although NO_x emissions were not actually subdivided according to the NO and NO_2 species.

In estimating uncertainties for individual source and source category emissions, the percentage uncertainties (precision values) associated with individual error sources were added linearly instead of quadratically, which would have been a more logical procedure to estimate the most probable error. When the uncertainties from individual sources and different source categories were combined, however, the conventional quadratic addition procedure was used.

The measure adopted by the U.S. group to express the degree of uncertainty was *probable error*, which was described as an estimate of the error in the data resulting from biases and imprecision. In determining the probable error for an emission factor derived from a sufficient amount of background data, a rigorous statistical analysis was performed to account for the error resulting from the assignment of an average emission factor to a particular source or group of sources (i.e., representation error). This analysis defined error as a function of:

1. The number of observations used to derive each factor,

2. The variability of those observations, and

3. The number of sources to which each factor might typically be applied in a statewide inventory.

For other EEPs, uncertainty estimates were assigned by source category to the state level based on engineering judgment or other simplistic assumptions.

The method used by the U.S. group to estimate uncertainties for emissions from individual source categories at the state level and for national total emissions conforms closely to the principles of quadratic addition.

A.5 NAPAP WORKSHOP REPORT

The NAPAP workshop report describes the uncertainty values assigned to EEPs at a workshop using a modified Delphi technique. The uncertainty values were developed for use in estimating the uncertainty associated with the 1980 NAPAP emissions inventories. Uncertainty estimates for the EEPs in Eq. A.1, temporal allocation factors, and

spatial disaggregation factors for area sources were made for SO_2, NO_x, and volatile organic compounds (VOC). Additional uncertainty estimates were assigned for NO_x and VOC speciation factors, and for non-1980 data in representing 1980 emissions. The degree of uncertainty for emission factors was expressed at the workshop by assuming that 90% of the values for an individual source lie within the mean uncertainty estimates.

The report states that a modified Delphi approach was used at the workshop, but it does not explicitly describe what the modification involved. Many assumptions were used to simplify the task. For example, it was assumed that the uncertainty estimates for most EEPs are the same for all SCCs.

APPENDIX B:

STATISTICAL FORMULAS FOR ESTIMATING AND AGGREGATING EEP UNCERTAINTIES

APPENDIX B:

STATISTICAL FORMULAS FOR ESTIMATING AND AGGREGATING EEP UNCERTAINTIES

B.1 DESCRIPTIVE MEASURES OF DATA VARIABILITY

For expressing the degree of data variability, one of the following four measures of dispersion or spread of a set of values $X_1, \ldots, X_n$ can be used:

1. Range R, or the difference between the minimum and maximum values of X_i,

2. Sample standard deviation S_n about the arithmetic mean $\bar{X}$, where

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 \qquad \text{(B.1)}$$

3. Relative standard deviation $S_n/\bar{X}$, also referred to as the coefficient of variation (CV), or

4. Precision of the mean (P), defined by:

$$P = \left(\frac{S}{\sqrt{n}}\right)\left(\frac{1}{\bar{X}}\right) \qquad \text{(B.2)}$$

The quantities R, S, CV, and P can easily be calculated for any arbitrary set of estimates or measurements without considering how those values are statistically distributed. These quantities are presented for selected EEPs on the data base forms in App. D.

B.2 ESTIMATING PARAMETER UNCERTAINTIES

The methods developed in this appendix for constructing confidence intervals for EEPs apply to both emission factors and temporal allocation factors. For presentation purposes, the methods are illustrated for emission factors only.

In order to make precise statements regarding emission factor uncertainty, assumptions about the process that generates the emission factor data are necessary. In this section, we restrict attention to the case where sampling variability is the only source of uncertainty. If there are other sources of error (e.g., systematic errors or a sample population that is not fully representative of the inventory population), then the confidence intervals that are derived will underestimate or provide a lower limit on the amount of uncertainty. The confidence intervals, especially when they imply very precise estimates of the parameters, should therefore be interpreted with caution.

Let $X_1, \ldots, X_n$ denote a sample of n individual-source emission factors from the population of all emission factors. Let μ be the average of the emission factors for all sources in the population. An unbiased estimator for μ is $\bar{X}_n$, the sample mean. Also, the variance of $\bar{X}_n$ is given by

$$\text{Var}(\bar{X}_n) = \frac{\sigma^2}{n} \tag{B.3}$$

where σ^2 is the variance of emission factors in the population.

To provide uncertainty estimates for $\bar{X}_n$ as an estimator for μ requires an assumption regarding the distribution of emission factors. In this appendix, we assume that this distribution is normal. This assumption greatly simplifies the analysis, though it may be unrealistic. One problem for further inquiry is the sensitivity of the analysis to the normality assumption. Another problem, which is discussed below, is estimating the uncertainty in the emission factor for a randomly drawn source (or a sample mean of a group of randomly drawn sources).

B.2.1 Confidence Intervals for a Parent Population Mean

Assume that the distribution of emission factors in the population is normal with unknown mean μ and unknown variance σ^2. Figure B.1a shows such a distribution. Let

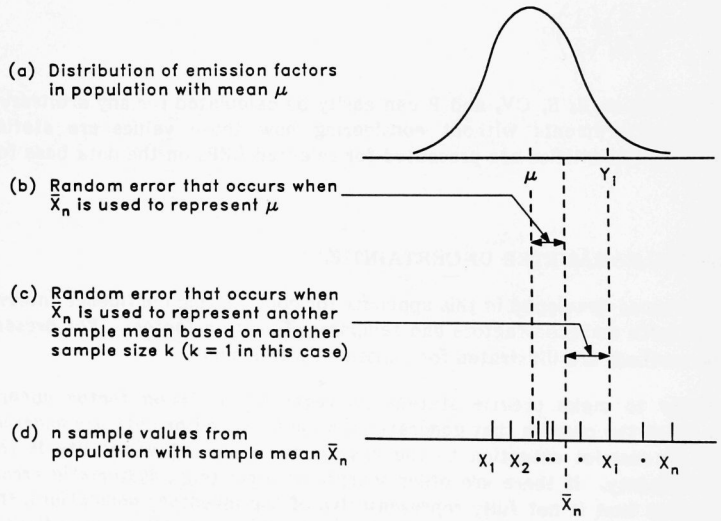


FIGURE B.1 Random Errors Due to Data Variability

$X_1, \ldots, X_n$ be a random sample from this population; an illustrative set of such values is presented in Fig. B.1d. If $\bar{X}_n$ denotes the sample mean and S_n^2 the sample variance as defined in Sec. B.1, then a $100(1 - \alpha)\%$ confidence interval for μ is given by

$$\bar{X}_n \pm t_{\alpha/2} \sqrt{s_n^2/n} \tag{B.4}$$

where $t_{\alpha/2}$ is the critical value of the Student's distribution with n - 1 degrees of freedom. The relative length of the interval, as a function of n, is given in Table B.1 in the $k = \infty$ column. (The k term appears in Eq. B.5, on which Table B.1 is based; however, it drops out when $k = \infty$, making Eqs. B.4 and B.5 identical.) For example, if n = 10, $S_n^2 = 1$, and $\bar{X}_n = 50$, then the length of the 95% confidence interval is 1.43 and its bounds, assuming normal distribution, are 49.28 and 50.72. The exact length of the interval and corresponding bounds are contingent on the assumptions regarding the distribution of the data.

In Tables 3.3, 3.4, and 3.7-3.11, Eq. B.4 is used to estimate the 95% confidence intervals for the population means of various EEPs. These confidence intervals are for the random errors that occur when the sample mean of data points is used to represent the mean of the parent population (see Fig. B.1b).

B.2.2 Confidence Intervals Applied to One or More Samples or Sources

Equation B.4 yields the $100(1 - \alpha)\%$ confidence interval when $\bar{X}_n$ is used to estimate μ. In this section we consider a variation -- the confidence interval when $\bar{X}_n$ is used to estimate the mean of another group from a population whose expected value is μ.

Suppose that the test sample on which the mean emission factor is based consists of n separately measured emission factors and that the sample mean is $\bar{X}_n$ and the sample variance is S_n^2. Suppose further that the group of interest is size k and that its mean is denoted by $\bar{Y}_k$. ($Y_i = \mu + \varepsilon_i$ is the true emission factor for the ith source and $\bar{Y}_k$ is the average over all k sources.) Assume that this group is independent of the test sample and that both have a common normal distribution. Then, the $100(1 - \alpha)\%$ confidence interval for $\bar{Y}_k$ is given by

$$\bar{X}_n \pm t_{\alpha/2} \sqrt{s_n^2 \left(\frac{1}{n} + \frac{1}{k}\right)} \tag{B.5}$$

where $t_{\alpha/2}$ is the critical value for the Student's distribution with n - 1 degrees of freedom.*

*The statistical meaning of the confidence interval given by formula B.5 can be described in terms of repeated sampling of both samples. Suppose a sample of X_i of size n and a sample of Y_i of size k are drawn and the confidence interval is calculated using formula B.5. The sample mean Y_k will fall within this interval with a frequency of $100(1 - \alpha)\%$.

To see why formula B.5 gives the correct interval, consider the random variable

$$(\bar{X}_n - \bar{Y}_k)\Big/\sqrt{s_n^2\left(\frac{1}{n} + \frac{1}{k}\right)} \tag{B.6}$$

and divide the numerator and denominator by the standard deviation of $\bar{X}_n - \bar{Y}_k$,

$$\sqrt{\sigma^2\left(\frac{1}{n} + \frac{1}{k}\right)}$$

where σ^2 is the variance of the population distribution. The numerator is then

$$(\bar{X}_n - \bar{Y}_k)\Big/\sqrt{\sigma^2\left(\frac{1}{n} + \frac{1}{k}\right)}$$

TABLE B.1 Relative Length of Confidence Intervals[a]

% Confidence Interval	n	Values of k						
		1	2	4	10	30	100	∞
90	2	15.5	12.6	10.9	9.78	9.22	9.00	8.92
95		31.1	25.4	22.0	19.70	18.61	18.00	18.00
90	4	5.26	4.08	3.33	2.78	2.50	2.40	2.35
95		7.12	5.51	4.50	3.77	3.39	3.24	3.18
90	10	3.84	2.84	2.17	1.64	1.34	1.21	1.16
95		4.74	3.50	2.68	2.02	1.65	1.49	1.43
90	30	3.45	2.48	1.81	1.24	0.88	0.71	0.62
95		4.16	2.99	2.18	1.49	1.06	0.86	0.74
90	∞	3.29	2.33	1.65	1.04	0.60	0.33	0
95		3.92	2.77	1.96	1.24	0.71	0.39	0

[a]The intervals are for $S_n^2 = 1$. The length of the confidence interval is defined as:

$$2\left[t_{\alpha/2}\sqrt{s_n^2\left(\frac{1}{n} + \frac{1}{k}\right)}\right]$$

where $t_{\alpha/2}$ is the critical value of the Student's distribution with $n - 1$ degrees of freedom.

which is the sum of normal random variables and is therefore itself normal with mean 0 and variance 1. The denominator is the square root of

$$\frac{(s_n^2)\left(\frac{1}{n} + \frac{1}{k}\right)}{(\sigma^2)\left(\frac{1}{n} + \frac{1}{k}\right)} = \frac{1}{n-1} \sum_{i=1}^{n} \left(\frac{X_i - \bar{X}_n}{\sigma}\right)^2$$

which is a chi-squared random variable with n - 1 degrees of freedom. Hence, the random variable given above (B.6) is a Student's t random variable with n - 1 degrees of freedom. Solving for the associated confidence interval yields formula B.5. With this formula, we can construct confidence intervals given data for $\bar{X}_n$, S_n^2, and k. The relative ranges of the intervals as a function of n, k, and α are given in Table B.1.

The 95% confidence interval based on formula B.5 is used in Tables 3.3, 3.4, and 3.7-3.11 to represent the errors that occur when a sample mean $\bar{X}_n$ is used to represent another sample mean based on a sample size of k (see Fig. B.1c).

B.2.3 Uncertainty with Systematic Bias

In this section, we briefly indicate the relationship between the uncertainty due to sampling variability and the uncertainty due to the possibility that the test sample may not be representative of the population of interest. The population indicated at the top of Figure B.2a is the one of interest. The actual emission factor at the ith source is $Y_i = \mu + \varepsilon_i$, where the frequency with which Y_i takes various values is illustrated in Fig. B.2a. The average emission factor over k sources is

$$\bar{Y}_k = \frac{1}{k} \sum_{i=1}^{k} Y_i$$

and is the quantity we want to estimate. To do so, we estimate $\bar{Y}_k$ by a sample mean from another independent set of observations, i.e., by

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

In this report, we have assumed that the independent set of observations has an identical distribution to Y_i so that the error of our estimate is solely due to sampling variability. While a complete discussion of the impacts of other sources of error is beyond the scope of this report, we want to indicate in a qualitative way how the methodology will change when the test sample may not be representative of the population of interest. Hence, we allow for the possibility of systematic bias by assuming that $X_i = \Delta + \mu + \varepsilon_i$. That is, X_i has the same frequency as Y_i except that it is shifted by an amount Δ. This is illustrated in Fig. B.2, along with a sample from the X_i distribution (Fig. B.2e). The figure illustrates, in a qualitative fashion, how our inferences about $\bar{Y}_k$ can be influenced by both sampling variability (Figs. B.1b and B.1c)

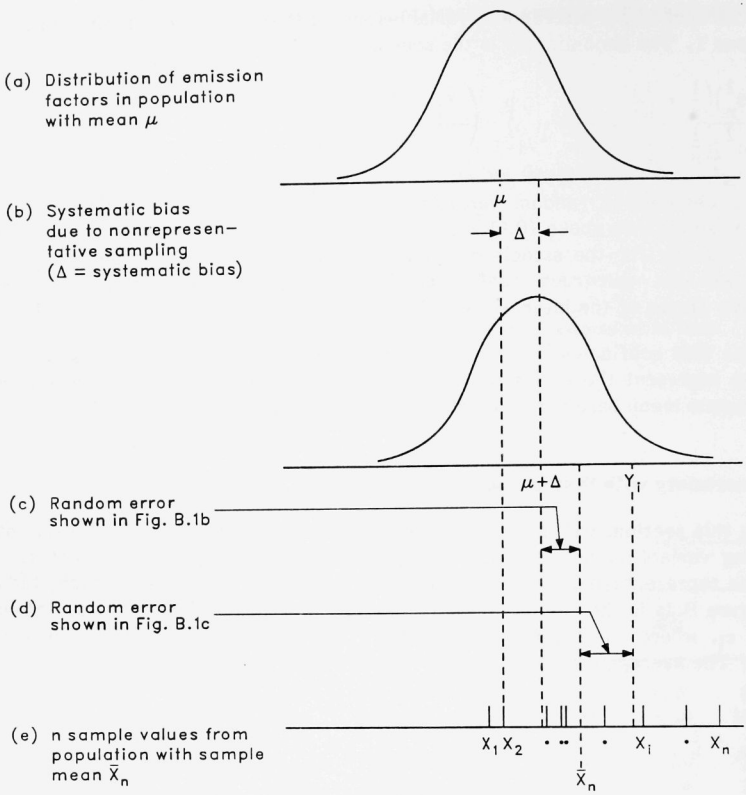


FIGURE B.2 Systematic Bias Due to Nonrepresentative Sampling and Random Errors Due to Data Variability

and systematic bias (Fig. B.1b). If systematic bias is unlikely, or if it is likely to be small in magnitude, then our uncertainty estimates will not be greatly affected.

B.3 FORMULAS FOF COMBINING UNCERTAINTY

Uncertainty estimates require the calculation of a measure of dispersion (as expressed usually by variance) of a random variable, which in some cases is a function of other random variables. In this section we present a variety of formulas for combining the uncertainty or dispersion of random variables together.

B.3.1 Variance of a Sum of Independent Random Variables

Let $X_1, \ldots, X_n$ be independent random variables where the variance of X_i is denoted by $\mathrm{Var}(X_i)$. (Actually, we only require the slightly weaker condition that the X_i values be uncorrelated.) Then,

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i) \tag{B.7}$$

This well-known equality is sometimes referred to as Beinayme's equality[1] and may be obtained as a simple exercise using the variance definition. The formula is useful for computing the variance of total emissions as the sum of individual emissions.

B.3.2 Variance of a Product of Independent Random Variables

Given the same conditions as in Sec. B.3.1, let the mean of X_i be μ. Then

$$\mathrm{Var}\left(\prod_{i=1}^{n} X_i\right) = \left(\prod_{i=1}^{n} \mu_i\right)^2 \left\{ \prod_{i=1}^{n} \left[1 + \frac{\mathrm{Var}(X_i)}{\mu_i^2}\right] - 1 \right\} \tag{B.8}$$

This formula is sometimes known as Goodman's formula,[2] and it is also derived in Benkovitz and Oden.[3] It may be used for computing the variance of emissions that have been calculated as the product of such terms as the emission factor, activity level, and temporal allocation factor. Unlike the sum of normal random variables, the product of such variables is not itself normally distributed. Thus, the standard procedures for estimating uncertainty intervals do not strictly apply to products of random variables.

B.3.3 Fractional Errors

For some applications, it has been proposed to replace the variance measure of dispersion with other measures. One such measure is the so-called fractional error considered by Benkovitz and Oden.[3] This measure is denoted by p(X) and is defined as

$$p(X) = \frac{z\sqrt{\mathrm{Var}(X)}}{\mu} \tag{B.9}$$

where z is an α percentage point of the distribution of X and $\mu \neq 0$. In applications where X is assumed normal and $\alpha = 0.95$, the value of $z = 1.96$. The advantage of p(X) as a dispersion or uncertainty measure is that it corresponds to questions often asked in opinion surveys about percentage errors. Its main disadvantage is the dependence on μ. The fractional error decreases when μ increases even though no corresponding change is needed in the shape of the X distribution.

The variance formulas given above can be used to derive corresponding formulas for p(X). For the fractional error of a sum, we have from the definition

$$p^2(\Sigma X_i) = \frac{z^2 \, \mathrm{Var}(\Sigma X_i)}{E^2(\Sigma X_i)} \tag{B.10}$$

Multiplying both sides by $E^2(\Sigma X_i)$ and using the variance of a sum formula then gives

$$\left[p^2(\Sigma X_i)\right]\left[E^2(\Sigma X_i)\right] = \sum \left\{ \left[p^2(X_i)\right]\left[\mu_i^2\right] \right\}$$

This formula, derived by Benkovitz and Oden,[3] is used in App. C to compute the fractional error of the sum of various terms.

In a similar fashion, the variance of a product formula may be used to obtain the fractional error of a product. We have

$$p^2(\Pi X_i) = \frac{z^2 \, \mathrm{Var}(\Pi X_i)}{E^2(\Pi X_i)}$$

which, by using Goodman's formula[3] and rearranging terms, yields

$$p^2(\Pi X_i) = (z^2)\left\{ \prod \left[1 + \frac{p^2(X_i)}{z^2} \right] - 1 \right\} \tag{B.11}$$

This formula, also derived by Benkovitz and Oden,[3] is used in App. C to compute the fractional error of the product of various terms.

B.4 DIRECTIONS FOR FUTURE WORK

The methods presented in this appendix are contingent on a number of simplifying assumptions, which may not be accurate. One area for future work is to accommodate more-realistic assumptions. Several specific topics are mentioned below.

First, it is possible that the test sample may not be representative of the population. This possibility will not alter our estimate of the average emission factor (unless there are reasons to suspect that the bias is in a certain direction), but it does contribute to uncertainty regarding the emission factor. One topic for future work is to develop methods for incorporating the possibility of systematic bias into the uncertainty estimates.

Second, one assumption used repeatedly is that the random variables follow a normal or Gaussian distribution. This assumption simplifies the analysis greatly but may not be representative of the data generation process and, most importantly, it can imply

confidence intervals that significantly underestimate the true uncertainty.[4] An investigation of uncertainty estimates under departures from the normality assumption is a topic for future work.

Finally, the sample mean used for the AP-42 emission factors may not be the best statistic for estimating the average emission factor in the source population. The sample mean is sensitive to extreme observed values (statistical outliers) and is typically the optimal statistic only in very restrictive situations that are not likely to be found in actual applications. This sensitivity has led to a variety of so-called "robust" statistics, which are nearly as good as the sample mean when the sample mean is optimal but which are far superior to the mean under a wide range of alternative and more realistic set of assumptions.[5] Assessing the usefulness of these robust procedures for emissions data is one more topic for future work.

B.5 REFERENCES

1. Loeve, M., *Probability Theory*, 3rd Ed., D. Von Nostrand, New York (1963).

2. Goodman, L.A., *On the Exact Variance of Products*, J. American Statistical Assn., 55:708-713 (1960).

3. Benkovitz, C.M., and N.L. Oden, *Methodology for Uncertainty Analysis of the NAPAP Emissions Inventory, Progress Report FY 1986*, draft, Brookhaven National Laboratory (Jan. 1987).

4. Hampel, F.R., et al., *Robust Statistics: The Approach Based on Influence Functions*, John Wiley and Sons, New York (1986).

5. Huber, P.J., *Robust Statistics*, John Wiley and Sons, New York (1981).

APPENDIX C:

METHODS AND ILLUSTRATION FOR ESTIMATING EMISSIONS UNCERTAINTIES

APPENDIX C:

METHODS AND ILLUSTRATION FOR ESTIMATING EMISSIONS UNCERTAINTIES

The methods and assumptions used in calculating uncertainties for the emissions estimates presented in Sec. 4 are described in this appendix, along with an example. The degree of uncertainty is expressed as the fractional error, i.e., the relative 95% confidence interval.

C.1 METHODS

The basic mathematical formula used in estimating uncertainties for emissions estimates for a single point source is Goodman's formula expressed in terms of variance:

$$\mathrm{Var}\left(\prod_{i=1}^{n} X_i\right) = \left(\prod_{i=1}^{n} \mu_i\right)^2 \left\{ \prod_{i=1}^{n} \left[1 + \frac{\mathrm{Var}(X_i)}{\mu_i^2} \right] - 1 \right\} \qquad \text{(C.1)}$$

or its equivalent expressed in terms of fractional error:

$$p^2\left(\prod_{i=1}^{n} X_i\right) = (z^2) \left\{ \prod_{i=1}^{n} \left[1 + \frac{p^2(X_i)}{z^2} \right] - 1 \right\} \qquad \text{(C.2)}$$

where:

$\mathrm{Var}(X_i)$ = variance of X_i,

μ_i = mean of X_i,

$\prod_{i=1}^{n} \mu_i$ = product of the means of X_i,

p = fractional error, and

z = a percentage point of the distribution of X_i and X, where X is the product of X_is.*

Equations C.1 and C.2 were discussed in App. B and correspond to Eqs. B.8 and B.11, respectively. These formulas allow computation of the variance or fractional error for an estimated emissions value, which is a product of several EEPs.

---

*If X is assumed to be normally distributed and $\alpha = 0.95$, then $z = 1.96$.

The relative 95% confidence interval for emissions is calculated as ±2 times the relative standard deviation, computed by using Goodman's formula. This procedure is consistent with the assumption that emissions are approximately normally distributed. Investigating emissions uncertainty assuming alternative distributions (e.g., lognormal) is one topic for future research.

Goodman's formula depends on the EEPs being uncorrelated. This assumption is sometimes quite inaccurate, because certain EEPs may be rather strongly correlated, whereas others are likely to be independent. Distinguishing these cases is another topic for future research.

For our example, annual and temporally allocated SO_x emissions for a single point source are estimated as follows:

- *Annual SO_x emissions (E_y)*:

$$E_y = (EF)(SC_y)(A_y) \quad \text{(C.3)}$$

where:

EF = AP-42 emission factor,

SC_y = annual mean fuel sulfur content, and

A_y = annual fuel throughput,

- *Specific-quarterly SO_x emissions (E_q)*, i.e., emissions for the first, second, third, or fourth quarter:

$$E_q = (EF)(SC_q)(A_y)(Q) \quad \text{(C.4)}$$

where:

SC_q = mean fuel sulfur content for the quarter, and

Q = quarterly activity allocation factor,

- *Mean hourly SO_x emissions (E_{mh})*, i.e., the mean for all of the specific hourly periods, e.g., 1-2 p.m., in a specific quarter:

$$E_{mh} = (EF)(SC_{mh})(A_y)(Q)(H) \quad \text{(C.5)}$$

where:

SC_{mh} = mean fuel sulfur content for all of the specific hourly periods, and

H = mean hourly activity allocation factor, and

- *Individual-hourly SO_x emissions (E_h)*, e.g., 1-2 p.m., April 25, 1985:

$$E_h = (EF)(SC_h)(A_y)(Q)(H)(c) \tag{C.6}$$

where:

SC_h = individual-hourly mean fuel sulfur content, and

c = individual-hourly activity correction factor for using a mean hourly activity value to represent the activity value for the same hour on a specific date (the expected value of c is 1).

The emissions equations for NO_x are identical to Eqs. C.3-C.6 except that there is no fuel sulfur content term.

The values of the EEPs in the emissions equations can be estimated in two different ways. In some cases, such as the annual activity level (fuel throughput) and annual mean fuel sulfur content, the values are assumed in this report to be estimated for each individual point source (i.e., these values are individual-source or unit values). In other cases, such as emission factors and temporal allocation factors, the values are estimated based on a small subset of sources in the relevant SCC category.

If the values of all EEPs were estimated for each individual source, as is possible in detailed field tests, the variance associated with the emissions from a single source could be computed using the variances for each EEP and Goodman's formula (Eq. C.1). The uncertainty values thus computed for individual-source emissions (assumed to be independent) can then be combined, using Beinayme's equality (in terms of variance), to calculate the uncertainty associated with emissions from multiple sources:

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i) \tag{C.7}$$

or its equivalent expressed in terms of fractional error:

$$\left[p^2(\Sigma X_i)\right]\left[E^2(\Sigma X_i)\right] = \sum \left\{\left[p^2(X_i)\right]\left[\mu_i^2\right]\right\} \tag{C.8}$$

Equations C.7 and C.8 were discussed in App. B and correspond to Eqs. B.7 and B.10, respectively.

However, for use in developing typical emissions inventories, most EEPs are not measured directly. Instead, they are represented by typical or mean values for the corresponding source category (which are referred to in this report as SCC-group mean estimates). When an SCC-group mean emission factor is used to represent more than one source, the emissions estimates from these sources are not independent (i.e., they are

based on the same AP-42 mean emission factor). In such cases, the above approaches are not accurate, because of this lack of independence. As indicated in Sec. 1, BNL is developing methods for computing uncertainties for emissions estimates based on the uncertainty values of various EEPs, including EEPs based on SCC-group mean estimates (a situation described by BNL as *subset sampling*). To date, a formula has been developed for a limited case involving only one EEP based on SCC-group mean estimates. Formulas for more-complicated cases, where two or more EEPs in the emissions equation are based on SCC-group mean values (e.g., both the emission factor and the temporal allocation factor), have yet to be developed.[1]

Goodman's formula provides an approximate solution for the variance of emissions when applied to k sources, where the variances of each EEP are calculated for k sources and combined as a product. We refer to this "top-down" approach, i.e., of calculating the EEP variances for k sources and then combining the results using Goodman's formula, as a *product-derived* approach. This approach is particularly applicable to area sources. Even in cases where Goodman's formula provides a quite accurate solution for the variance, the confidence interval constructed as ±2 times the standard deviation will not be exact, in general, since the distribution of emissions is unlikely to be exactly a normal distribution.

In the absence of an accurate methodology for general cases, Goodman's formula is used as an approximation to illustrate how the variances for various EEPs (including those based on either individual-source data or SCC-group mean estimates) may be combined to estimate an uncertainty value for emissions.

For the case of 100 point sources, the annual and temporally allocated emissions of SO_x are estimated as follows:

- *Annual SO_x emissions (E_{y100})*:

$$E_{y100} = (EF)(\overline{SC}_{y100})(A_{y100}) \qquad \text{(C.9)}$$

where:

$\overline{SC}_{y100}$ = annual mean fuel sulfur content for the 100 point sources, and

A_{y100} = total annual fuel throughput for the 100 point sources,

- *Specific-quarterly SO_x emissions (E_{q100})*:

$$E_{q100} = (EF)(Q)(\overline{SC}_{q100})(A_{y100}) \qquad \text{(C.10)}$$

where:

$\overline{SC}_{q100}$ = mean fuel sulfur content for the 100 point sources for the quarter,

- *Mean hourly SO_x emissions ($E_{mh}100$),*

$$E_{mh100} = (EF)(Q)(H)(\overline{SC}_{mh100})(A_{y100}) \quad \text{(C.11)}$$

where:

$\overline{SC}_{mh100}$ = mean fuel sulfur content for the 100 point sources for the same hourly period, and

- *Individual-hourly SO_x emissions (E_h100):*

$$E_h100 = (EF)(Q)(H)(\overline{SC}_{h100})(\overline{c}_{100})(A_{y100}) \quad \text{(C.12)}$$

where:

$\overline{SC}_{h100}$ = mean fuel sulfur content for the 100 point sources for that hourly period, and

$\overline{c}_{100}$ = mean individual-hourly activity correction factor for the 100 point sources.

The corresponding emissions equations for NO_x are identical to Eqs. C.9-C.12 except that there is no sulfur content term.

The variance for the product of the mean sulfur content, the mean individual-hourly activity correction factor, and the sum of annual activity levels can be computed using Goodman's formula. However, further computations cannot proceed for cases involving two or more group mean EEPs (i.e., based on subset sampling) because exact formulas to calculate the uncertainty values for emissions estimates for such cases are not currently available. To circumvent this problem, the group mean EEPs in Eqs. C.9-C.12, i.e., EF, Q, and H, are viewed as the mean values for the 100 sources, i.e., $\overline{EF}_{100}$, $\overline{Q}_{100}$, and $\overline{H}_{100}$.

With this, further computations can be carried out using Goodman's formula (Eq. C.1) or its equivalent (Eq. C.2) and the uncertainty values developed in Sec. 3 for the mean values for multiple (k) sources where k = 100. (See Tables 3.7, 3.9, and 3.10, and column 5 in Table C.1.) This is equivalent to the product-derived approach described in Sec. 1 for estimating total emissions for multiple sources, in which the total emissions are obtained by multiplying relevant EEPs estimated for multiple sources.

For comparison, an alternate approach for estimating the uncertainty associated with multiple-source emissions was considered. In this approach, uncertainty values estimated for a single source are combined for 100 statistically identical sources using Beinayme's equality (Eq. C.7) or its equivalent (Eq. C.8), as if all the source emission estimates were independent (which they are not, due to subset sampling for some of the EEPs). This is known as the "bottom-up" approach (which we call the *aggregate-derived* approach), which can be mathematically expressed as follows:

$$E_{y100} = \sum_{i=1}^{100} E_y \qquad \text{(C.13)}$$

$$E_{q100} = \sum_{i=1}^{100} E_q \qquad \text{(C.14)}$$

$$E_{mh100} = \sum_{i=1}^{100} E_{mh} \qquad \text{(C.15)}$$

$$E_{h100} = \sum_{i=1}^{100} E_h \qquad \text{(C.16)}$$

For the case of 100 identical emission sources, the equation for calculating the uncertainty for total emissions, Eq. C.8 (Beinayme's equality) reduces to

$$p^2(\Sigma X_i) = (1/100)\left[p^2(X_i)\right]$$

or

$$p(\Sigma X_i) = (1/10)\left[p(X_i)\right] \qquad \text{(C.17)}$$

C.2 ILLUSTRATION

This illustration covers annual and temporally allocated emissions for the following two cases:

1. One 500-MWe coal-fired electric utility boiler unit, burning raw coal and without SO_x scrubbing, and

2. 100 identical such units.

Uncertainty values (in terms of relative 95% confidence intervals) for the EEPs involved in this example are listed in Table C.1. With these EEP uncertainty values and

TABLE C.1 Estimated Uncertainty Values of Various EEPs for 500-MWe Coal-Fired Electric Utility Boiler Units

EEP (col. 1)	Relative 95% Confidence Interval (±%) For 1 Unit (col. 2)	For 100 Units (col. 3)	For the Mean EEP for k Units k = 1 (col. 4)	k = 100 (col. 5)
Annual activity level	0.1[a]	0.01[b]	NA	NA
Fuel sulfur content				
Annual[c]	0.6	0.06	NA	NA
Quarterly[c]	1.2	0.12	NA	NA
Hourly				
Mean[c]	5.9	0.59	NA	NA
Individual	41[d]	5.5[c]	NA	
AP-42 emission factor				
SO_x[e]	NA	NA	41	7
NO_x[f]	NA	NA	70	15
Temporal allocation factor				
Quarterly[g]	NA	NA	90	15
Mean hourly[h]	NA	NA	29	5
Individual hourly correction factor	170[i]	17[j]	NA	NA

[a]Data from Table 3.1 for the electric utility industry.

[b]Computed using Beinayme's equality (Eq. C.8).

[c]Based on the data in Table 3.3 for a 500-MWe unit burning raw coal and the assumption that the relative 95% confidence interval is ±2 times the relative standard deviation.

[d]Derived from the mean value for 91 hourly periods (e.g., the total number of 1-2 p.m. periods in a quarter).

[e]Values shown are those from Table 3.7 for all pulverized coal, spreader, and overfeed stoker boilers burning bituminous coal, 1-unit and 100-unit cases.

[f]Values shown are averages of those given in Table 3.7 for dry bottom and tangential boilers, 1-unit and 100-unit cases.

[g]Values shown are the mean values, for the 1-unit and 100-unit cases, of the four values given in Table 3.9 for each case.

[h]Values shown are the weighted averages, for the 1-unit and 100-unit cases, of the three values given in Table 3.10 for each case in the average variability category.

[i]Weighted average of the three values shown in Table 3.11 for the 1-hr mean in the average variability case.

[j]Computed, based on the value for one unit (col. 2).

NA = not applicable.

with Eqs. C.2 and C.17, the steps for computing emissions uncertainties are shown below. Since a number of strong assumptions are made and an approximate methodology is employed, the uncertainty values estimated below should not be considered as accurate estimates, but as rough approximations. These uncertainty estimates illustrate the usefulness of the EEP uncertainty values developed in Sec. 3.

C.2.1 Annual Emissions for a Single Unit

Annual SO_x emissions for one unit, E_y, are estimated using Eq. C.3:

$$E_y = (EF)(SC_y)(A_y) \tag{C.3}$$

The fractional errors (in terms of the relative 95% confidence interval) from Table C.1 for the EEPs in Eq. C.3 are as follows:

- Emission factor for SO_x: ±0.41
- Emission factor for NO_x: ±0.70
- Annual fuel sulfur content: ±0.006
- Annual activity level: ±0.001

The fractional error for annual SO_x emissions, calculated using Eq. C.2 (and assuming z = 2), is:

$$2\left\{\left[1 + (0.41/2)^2\right]\left[1 + (0.006/2)^2\right]\left[1 + (0.001/2)^2\right] - 1\right\}^{0.5}$$

$$= \pm 0.41$$

and the relative 95% confidence interval is ±41%. For annual NO_x emissions from one unit, the fractional error and relative 95% confidence interval are ±0.70 and ±70%, respectively.

C.2.2 Quarterly Emissions for a Single Unit

Annual SO_x emissions for one unit, E_q, are estimated using Eq. C.4:

$$E_q = (EF)(SC_q)(A_y)(Q) \tag{C.4}$$

The fractional errors from Table C.1 for the EEPs in Eq. C.4 are as follows:

- Emission factor for SO_x: ±0.41
- Emission factor for NO_x: ±0.70

- Quarterly fuel sulfur content: ±0.012
- Annual activity level: ±0.001
- Quarterly allocation factor: ±0.90

The fractional error for quarterly SO_x emissions, calculated using Eq. C.2 (and assuming z = 2), is:

$$2\left\{\left[1 + (0.41/2)^2\right]\left[1 + (0.012/2)^2\right]\left[1 + (0.001/2)^2\right]\left[1 + (0.90/2)^2\right] - 1\right\}^{0.5} = \pm 1.01$$

and the relative 95% confidence interval is ±101%. For quarterly NO_x emissions from one unit, the fractional error and relative 95% confidence interval are ±1.18 and ±118%, respectively.

C.2.3 Mean Hourly Emissions for a Single Unit

Mean hourly SO_x emissions for one unit, E_{mh}, are estimated using Eq. C.5:

$$E_{mh} = (EF)(SC_{mh})(A_y)(Q)(H) \qquad \text{(C.5)}$$

The fractional errors from Table C.1 for the EEPs in Eq. C.5 are as follows:

- Emission factor for SO_x: ±0.41
- Emission factor for NO_x: ±0.70
- Mean hourly fuel sulfur content: ±0.059
- Annual activity level: ±0.001
- Quarterly allocation factor: ±0.90
- Mean hourly allocation factor: ±0.29

The fractional error for mean hourly SO_x emissions, calculated using Eq. C.2 (and assuming z = 2), is:

$$2\left\{\left[1 + (0.41/2)^2\right]\left[1 + (0.059/2)^2\right]\left[1 + (0.001/2)^2\right]\left[1 + (0.90/2)^2\right]\left[1 + (0.29/2)^2\right] - 1\right\}^{0.5} = \pm 1.06$$

and the relative 95% confidence interval is ±106%. For NO_x emissions from one unit, the fractional error and relative 95% confidence interval are ±1.23 and ±123%, respectively.

C.2.4 Individual-Hourly Emissions for a Single Unit

Individual-hourly emissions for one unit, E_h, are estimated using Eq. C.6:

$$E_h = (EF)(SC_h)(A_y)(Q)(H)(c) \qquad (C.6)$$

The fractional errors from Table C.1 for the EEPs in Eq. C.6 are as follows:

- Emission factor for SO_x: ±0.41
- Emission factor for NO_x: ±0.70
- Individual-hourly fuel sulfur content: ±0.41
- Annual activity level: ±0.001
- Quarterly allocation factor: ±0.90
- Mean hourly allocation factor: ±0.29
- Individual-hourly correction factor: ±1.70

The fractional error for individual-hourly SO_x emissions, calculated using Eq. C.2 (and assuming z = 2), is:

$$2\left\{\left[1 + (0.41/2)^2\right]\left[1 + (0.41/2)^2\right]\left[1 + (0.001/2)^2\right]\left[1 + (0.90/2)^2\right]\left[1 + (0.29/2)^2\right]\left[1 + (1.70/2)^2\right] - 1\right\}^{0.5} = \pm 2.28$$

and the relative 95% confidence interval is ±228%. For NO_x emissions from one unit, the fractional error and relative 95% confidence interval are ±2.34 and ±234%, respectively.

C.2.5 Annual Emissions for 100 Identical Units

We have shown earlier that annual SO_x emissions from 100 identical units can be estimated according to Eq. C.9 (with EF replaced by $\overline{EF}_{100}$) or Eq. C.13:

$$E_{y100} = (\overline{EF}_{100})(\overline{SC}_{y100})(A_{y100}) \qquad (C.9)$$

or

$$E_{y100} = \sum_{i=1}^{100} E_y \qquad \text{(C.13)}$$

The fractional errors from Table C.1 for the EEPs in Eq. C.9 are as follows:

- Emission factor for SO_x: ±0.07
- Emission factor for NO_x: ±0.15
- Annual fuel sulfur content: ±0.0006
- Annual activity level: ±0.0001

The fractional errors for annual SO_x and NO_x emissions from one unit are ±0.41 and ±0.70, respectively, as calculated in Sec. C.2.1.

The fractional error for annual SO_x emissions from 100 units, based on the product-derived approach (Eq. C.9), is calculated using Eq. C.2 as:

$$2\left\{\left[1 + (0.07/2)^2\right]\left[1 + (0.0006/2)^2\right]\left[1 + (0.0001/2)^2\right] - 1\right\}^{0.5}$$

$$= \pm 0.07$$

and the relative 95% confidence interval is ±7%. For NO_x emissions from 100 units, the fractional error and relative 95% confidence interval are ±0.15 and ±15%, respectively.

The fractional errors for annual SO_x and NO_x emissions from 100 units, based on the aggregate-derived approach (Eq. C.13), are simply 1/10 of those for one unit, or ±0.04 for SO_x and ±0.07 for NO_x. The corresponding relative 95% confidence intervals in the 100-unit case are ±4% for SO_x and ±7% for NO_x.

C.2.6 Quarterly Emissions for 100 Identical Units

The two ways of computing the estimated quarterly SO_x emissions for 100 identical units are Eqs. C.10 (with EF and Q replaced by $\overline{EF}_{100}$ and $\bar{Q}_{100}$, respectively) and C.14:

$$E_{q100} = (\overline{EF}_{100})(\bar{Q}_{100})(\overline{SC}_{q100})(A_{y100}) \qquad \text{(C.10)}$$

and

$$E_{q100} = \sum_{i=1}^{100} E_q \qquad \text{(C.14)}$$

The fractional errors from Table C.1 for the EEPs in Eq. C.10 are as follows:

- Emission factor for SO_x: ±0.07
- Emission factor for NO_x: ±0.15
- Quarterly allocation factor: ±0.15
- Quarterly fuel sulfur content: ±0.0012
- Annual activity level: ±0.0001

The fractional errors for quarterly SO_x and NO_x emissions for one unit are ±1.01 and ±1.18, respectively, as calculated in Sec. C.2.2.

The fractional error for quarterly SO_x emissions from 100 units, based on the product-derived approach (Eq. C.10), is calculated using Eq. C.2 as:

$$2\left\{\left[1 + (0.07/2)^2\right]\left[1 + (0.15/2)^2\right]\left[1 + (0.0012/2)^2\right]\left[1 + (0.0001/2)^2\right] - 1\right\}^{0.5} = \pm 0.17$$

and the relative 95% confidence interval is ±17%. For NO_x emissions from 100 units, the fractional error and relative 95% confidence interval are ±0.21 and ±21%, respectively.

The fractional errors for quarterly SO_x and NO_x emissions from 100 units, based on the aggregate-derived approach (Eq. C.14), are simply 1/10 of those for one unit, or ±0.10 for SO_x and ±0.12 for NO_x. The corresponding relative 95% confidence intervals in the 100-unit case are ±10% for SO_x and ±12% for NO_x.

C.2.7 Mean Hourly Emissions for 100 Identical Units

The two ways of computing the estimated mean hourly SO_x emissions for 100 identical units are Eqs. C.11 (with EF, Q, and H replaced by $\overline{EF}_{100}$, $\bar{Q}_{100}$, and $\bar{H}_{100}$, respectively) and C.15:

$$E_{mh100} = (\overline{EF}_{100})(\bar{Q}_{100})(\bar{H}_{100})(\overline{SC}_{mh100})(A_{y100}) \qquad (C.11)$$

and

$$E_{mh100} = \sum_{i=1}^{100} E_{mh} \qquad (C.15)$$

The fractional errors from Table C.1 for the EEPs in Eq. C.11 are as follows:

- Emission factor for SO_x: ±0.07
- Emission factor for NO_x: ±0.15
- Quarterly allocation factor: ±0.15
- Mean hourly allocation factor: ±0.05
- Mean hourly fuel sulfur content: ±0.0059
- Annual activity level: ±0.0001

The fractional errors for mean hourly SO_x and NO_x emissions from one unit are ±1.06 and ±1.23, respectively, as calculated in Sec. C.2.3.

The fractional error for mean hourly SO_x emissions from 100 units, based on the product-derived approach (Eq. C.11), is calculated using Eq. C.2 as:

$$2\left\{\left[1 + (0.07/2)^2\right]\left[1 + (0.15/2)^2\right]\left[1 + (0.05/2)^2\right]\left[1 + (0.0059/2)^2\right]\left[1 + (0.0001/2)^2\right] - 1\right\}^{0.5} = \pm 0.17$$

and the relative 95% confidence interval is ±17%. For NO_x emissions from 100 units, the fractional error and relative 95% confidence interval are ±0.22 and ±22%, respectively.

The fractional errors for mean hourly SO_x and NO_x emissions from 100 units, based on the aggregate-derived approach (Eq. C.15), are 1/10 of those for one unit, or ±0.11 for SO_x and ±0.12 for NO_x. The corresponding relative 95% confidence intervals in the 100-unit case are ±11% for SO_x and ±12% for NO_x.

C.2.8 Individual-Hourly Emissions for 100 Identical Units

The two ways of computing the estimated individual-hourly SO_x emissions for 100 identical units are Eqs. C.12 (with EF, Q, and H replaced by EF_{100}, $\bar{Q}_{100}$, and $\bar{H}_{100}$, respectively) and C.16:

$$E_{h100} = (\overline{EF}_{100})(\bar{Q}_{100})(\bar{H}_{100})(\overline{SC}_{h100})(\bar{c}_{100})(A_{y100}) \qquad \text{(C.12)}$$

and

$$E_{h100} = \sum_{i=1}^{100} E_h \qquad \text{(C.16)}$$

The fractional errors from Table C.1 for the EEPs in Eq. C.12 are as follows:

- Emission factor for SO_x: ±0.07
- Emission factor for NO_x: ±0.15
- Quarterly allocation factor: ±0.15
- Mean hourly allocation factor: ±0.05
- Individual-hourly fuel sulfur content: ±0.055
- Individual-hourly correction factor: ±0.17
- Annual activity level: ±0.0001

The fractional errors for individual-hourly SO_x and NO_x emissions from one unit are ±2.28 and ±2.34, respectively, as calculated in Sec. C.2.4.

The fractional error for individual-hourly SO_x emissions from 100 units, based on the product-derived approach (Eq. C.12), is calculated using Eq. C.2 as:

$$2\left\{\left[1 + (0.07/2)^2\right]\left[1 + (0.15/2)^2\right]\left[1 + (0.05/2)^2\right]\left[1 + (0.055/2)^2\right]\right.$$
$$\left.\left[1 + (0.17/2)^2\right]\left[1 + (0.0001/2)^2\right] - 1\right\}^{0.5} = \pm 0.25$$

and the relative 95% confidence interval is ±25%. For NO_x emissions from 100 units, the fractional error and relative 95% confidence interval are ±0.28 and ±28%, respectively.

The fractional errors for individual-hourly SO_x and NO_x emissions from 100 units, based on the aggregate-derived approach (Eq. C.8), are simply 1/10 of those for one unit, or ±0.23 for both SO_x and NO_x. The corresponding relative 95% confidence intervals in the 100-unit case are ±23% for both SO_x and NO_x.

C.2.9 Summary

Table C.2 summarizes all of the uncertainty values estimated in Secs. C.2.1-C.2.8 for annual and temporally allocated SO_x and NO_x emissions estimates for a single unit and a group of 100 identical boiler units. The fractional uncertainty values calculated are listed in terms of percentages after multiplying by 100%. The bounds of uncertainty estimates for the single unit's emissions are listed in column 2, and those for the 100 units' emissions are listed in column 3 (product-derived approach) and column 4 (aggregate-derived approach).

The lower bounds for the relative 95% confidence are truncated at -100% in 6 of the 16 cases illustrated in Table C.2, and all 6 of these cases involve quarterly or hourly estimates for a single unit. These cases illustrate that it is difficult to accurately

TABLE C.2 Uncertainty Estimates for SO_x and NO_x Emissions from One or a Group of 100 Identical Coal-Fired Electric Utility Boiler Units[a]

Pollutant, Period (col. 1)	Relative 95% Confidence Interval[b] (±%)		
		Emissions from 100 Units	
	Single-Unit Emissions (col. 2)	Product-Derived Approach (col. 3)	Aggregate-Derived Approach[c] (col. 4)
SO_x			
Year	41	7	4
Quarter	101[d]	17	10
Hour			
Mean	106[d]	17	11
Individual	228[d]	25	23
NO_x			
Year	70	15	7
Quarter	118[d]	21	12
Hour			
Mean	123[d]	22	12
Individual	234[d]	28	23

[a]No emission control is assumed.

[b]Expressed as ±% of the emissions estimate for each period.

[c]This approach neglects the uncertainty due to an insufficient amount of measurement data used in developing mean EEP values.

[d]Confidence intervals that extend beyond -100% should be truncated at -100%.

estimate temporally allocated emissions for a single unit. The practical implication of the need to truncate the preliminary and illustrative results is that a method must be developed for performing calculations that involve variables that are not normally distributed. In those cases, the reasonableness of the upper bounds must also be called into question.

There are several alternatives to the normal distribution that should be investigated in the future. Among these are the lognormal, beta, and normal-on-log-odds (NOLO) distributions, all of which are primarily skewed distributions. The lognormal distribution has the attractive property that the product of lognormally distributed variables is also lognormal, with a mean and variance that are easily calculated.[2] Beta[3] and NOLO[4] distributions can easily be constructed over specific intervals. Lognormal distributions can also be defined over specific intervals, but the mathematics is more difficult. If the mean is sufficiently far away from a natural boundary (such as negative numbers), then the lognormal distribution can be used to obtain good results without restricting the variables to specific intervals. If all of the variables in an emissions equation can be approximated with lognormal distributions, then the mathematical computations are straightforward. However, if a number of distributions are needed, other methods must be developed to calculate the desired results.

C.3 REFERENCES

1. Benkovitz, C.M., and N.L. Oden, *Methodology for Uncertainty Analysis of the NAPAP Emissions Inventory, Progress Report FY 1986*, draft, Brookhaven National Laboratory (Jan. 1987).

2. Aitchison, J., and J.A.C. Brown, *The Lognormal Distribution: with Special Reference to Its Uses in Economics*, Cambridge University Press, London (1957).

3. Winkler, R.L., *An Introduction to Bayesian Inference and Decision*, Holt, Rinehart, and Winston, New York (1972).

4. Whitfield, R.G., and T.S. Wallsten, *Estimating Risks of Lead-Induced Hemoglobin Decrements under Conditions of Uncertainty: Methodology, Pilot Judgments, and Illustrative Calculations*, Argonne National Laboratory Report ANL/EES-TM-276 (Sept. 1984).

APPENDIX D:

EEP VARIABILITY DATA BASE

APPENDIX D:

EEP VARIABILITY DATA BASE

D.1 DATA BASE FORMS

This data base contains completed data base forms, as described in Sec. 3.1, for selected EEPs. For each EEP, the forms provide the mean, number of data points on which the mean was based, extreme values, standard deviation, CV values, and relative extreme values. Specifically, the following EEPs are covered:

- FGD system penetration factors (Tables D.1 and D.2),
- SO_x emission factors (Tables D.3-D.8), and
- NO_x emission factors (Tables D.9-D.26).

Section D.2 lists the data sources cited on the forms.

D.2 REFERENCES

1. *Source Inventory and Emission Factor Analysis, Vols. 1 and 2,* prepared by PEDCo-Environmental Specialists, Inc., U.S. Environmental Protection Agency Report EPA-450/3-75-082-a,b (Sept. 1974).

2. Peduto, E.F., Jr., et al., *Statistical Analysis of Continuous SO_2 Emissions Data from Steam Generators: Vol. I, A.B. Brown Unit No. 1; Vol. II, Springfield City Utilities, Southwest Unit No. 1; Vol. III, Louisville Gas and Electric Company, Cane Run Unit No. 6;* and *Vol. IV, Allegheny Power Systems, Pleasants Unit No. 1,* prepared by GCA Corp. for U.S. Environmental Protection Agency, under Contract 68-02-3168, Task 53 (July 1983).

3. Smith, W.S., and C.W. Gruber, *Atmospheric Emissions from Coal Combustion — An Inventory Guide,* U.S. Public Health Service Report 999-AP-24 (1966).

4. Lahre, T., U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, N.C., personal communication (April 24, 1986).

5. Cuffe, S.T., et al., *Air Pollution Emissions from Coal-Fired Power Plants; Report No. 1,* J. of Air Pollution Control Assn., *14*:353-362 (1964).

6. Gerstle, R.W., et al., *Air Pollution Emissions from Coal-Fired Power Plants, Report No. 2,* J. of Air Pollution Control Assn., *15*:59-64 (1965).

7. Smith, W.S., *Atmospheric Emissions from Fuel Oil Combustion: An Inventory Guide,* U.S. Public Health Service Report 999-AP-2 (1962).

8. *Environmental Control Implications of Generating Electric Power from Coal; 1977 Technology Status Report, Appendix D: Assessment of NO_x Control Technology for Coal Fired Utility Boilers*, prepared by KVB Inc., Argonne National Laboratory Report ANL/ECT-3 (Dec. 1977).

TABLE D.1 Completed Data Base Form for FGD System Penetration Factors: Hourly Data

Min. 1.0S 1.0S Max. — Emissions-Estimation Parameter Variability Data Base — Low REV 1.0CV 1.0CV High REV

Date: 4/29/86

Parameter	FGD Penetration Factor, Species: SOx, Averaging Period: Hour
SCC Number	1-01-002-01-05
SCC Description	Ext. Comb. Boiler, Elec. Gen. Bit. Coal, Pulv.

Data Source	Mean ($\bar{x}$), Sample Standard Deviation (S), and Max. and Min. Values (%)	Sample Number
Ref. 2 a-1	1.0, 7.1, 11.8, 15.9	1352
a-2	10.1, 14.8, 24.1	1246
b-1	1.2, 3.3, 3.5, 13.5, 18.7, 28.8	720
b-2	1.5, 11.0, 16.1, 39.3	767
b-3	1.2, 5.2, 8.8, 19.4, 34.4	769
B-4	0.5, 5.7, 12.3, 20.3, 36.2	755
C-1	1.4, 1.4, 6.4, 13.0, 37.8	646
C-2	0.0, 2.8, 3.9, 9.0, 14.4	526
D-1	0.0, 1.3, 4.1, 5.9, 7.2, 10.3, 24.5, 26.3	2483
E-1	1.9, 9.7	1632
E-2	0.7, 0.9, 4.4, 5.8, 11.0, 17.6, 31.9, 32.2	1123

Data Source	Coefficient of Variation (CV), % and Relative Extreme Value (REV), %	
Ref. 2 a-1	-92, 35, 104	
a-2	-92, 25	
b-1	-76, 70, 113	
b-2	-89, 83, 257	291
b-3	-86, 58, 194	
b-4	-96, 56	
C-1	-89, 64, 191	269
C-2	-100, 53	315
D-1	-100, 43	265
E-1	-82, 67	450
E-2	-88, 60, 193	

-200 -100 0 100 200

-92

Comments:

a Dual alkali (FMC); b Limestone: Adipic Acid Enhanced: Case 1;
c Limestone: Adipic Acid Enhanced: Case 2;
d Dual Alkali (CEA/ADL); e Limestone: MgO promoted (AAF)

TABLE D.2 Completed Data Base Form for FGD System Penetration Factors: Daily Data

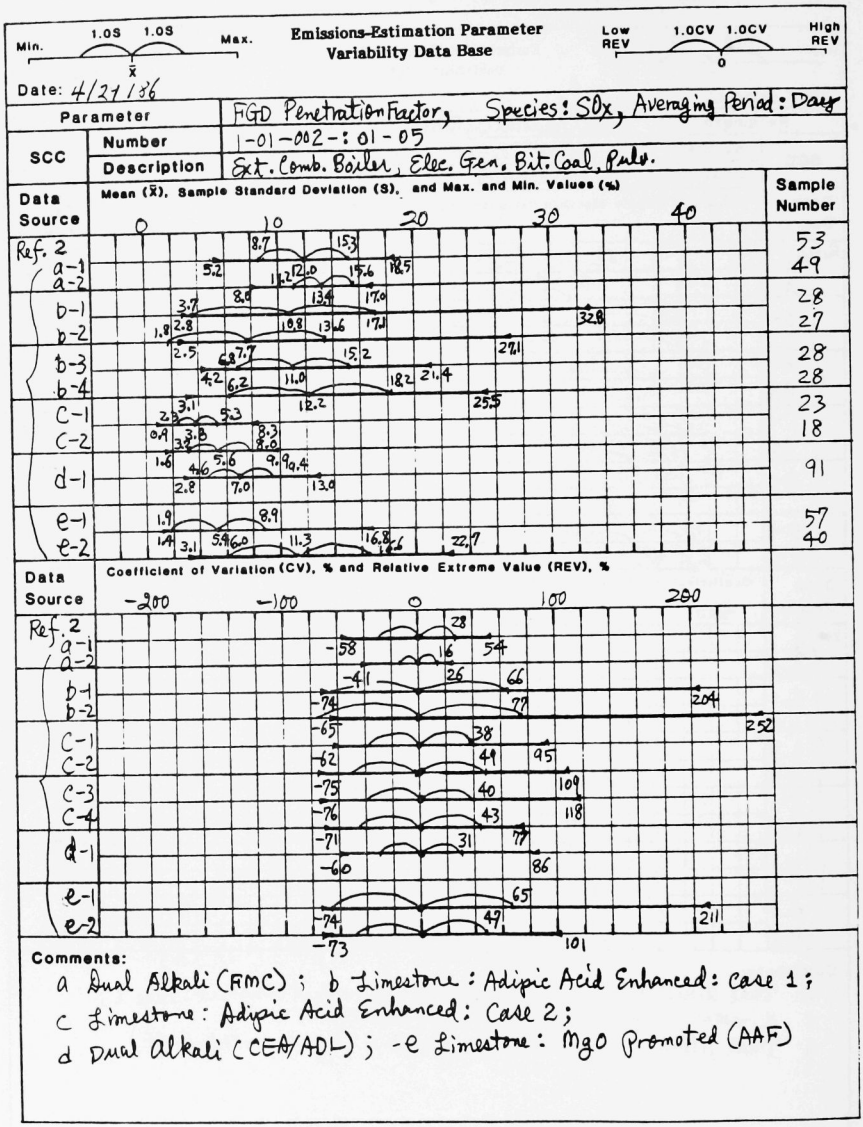


TABLE D.3 Completed Data Base Form for Emission Factors: SO_x Data, Form 1

Min. 1.0S 1.0S Max. (x̄)

Emissions-Estimation Parameter Variability Data Base

Low REV 1.0CV 1.0CV High REV (0)

Date: 4/9/86

Parameter		Emission Factor	Species: SOx
SCC	Number	1-01-002-: 01, 02, 03	
	Description	Ext. Comb. Boiler, Elec. Gen. Bit. Coal, Pulv.: Wet, Dry, Cyclone	

Data Source — Mean (x̄), Sample Standard Deviation (S), and Max. and Min. Values (lb of SO_2/ton) — Sample Number

20S[b] 30S 40S 50S 60S

Data Source	Mean (x̄), Sample Standard Deviation (S), and Max. and Min. Values (lb of SO_2/ton)	Sample Number
AP-42 (73 EF 82)	38S; 39S	
Ref. 1 Wet	28.7S; 24.5S; 38.3S; 47.9S; 56.9S	7
Dry	29.5S; 24.5S; 38.3S; 47.2S; 56.9S	8
Cyclone	33.5S; 31.9S; 36.7S; 39.9S; 38.7S	4

Data Source — Coefficient of Variation (CV), % and Relative Extreme Value (REV), %

−100 −50 0 50 100

Data Source	Coefficient of Variation (CV), % and Relative Extreme Value (REV), %
Ref. 1 Wet	25; −36; 49
Dry	23; −36; 49
Cyclone	9; −14; 5

Comments: a Expressed as SO_2, b Coal sulfur content in weight %.

TABLE D.4 Completed Data Base Form for Emission Factors: SO_x Data, Form 2

Min. 1.0S 1.0S Max. (x̄) — Emissions-Estimation Parameter Variability Data Base — Low REV 1.0CV 1.0CV High REV (0)

Date: 4/9/86

Parameter		Emission Factor — Species: SOx
SCC	Number	1-01-001, 002; 1-02-001, 002; 1-03-001, 002; 1-05-001-02
	Description	Coal Combustion (Domestic combustion & Electric Power Plant Boilers)

Data Source	Mean (x̄), Sample Standard Deviation (S), and Max. and Min. Values (lb of SO_2/ton)	Sample Number
	10S 20S 30S 40S 50S	
AP-42 73 EF 82	38.8	
	39.8	
Ref. 3	27.2S 39.5S	152
	13S 33.3S 41	

Data Source	Coefficient of Variation (CV), % and Relative Extreme Value (REV), %	
	−100 −50 0 50 100	
Ref. 3	18	
	-61 23	

Comments:

TABLE D.5 Completed Data Base Form for Emission Factors: SO_x Data, Form 3

Min. 1.0S 1.0S Max. (x̄)

Emissions-Estimation Parameter Variability Data Base

Low REV 1.0CV 1.0CV High REV (0)

Date: 5/2/86

Parameter		Emission Factor, Species: SO_x
SCC	Number	1-01-002-: 01-05, 12, 17, 21-26; 1-02-002-: 01-05, 10, 12, 17, 19, 21-29; 1-03-002-: 05-07, 09, 11, 16, 17, 21-26.
	Description	Ext. Comb. Boiler, Elec. Gen., Ind., Comm-Inst., Bit. & Subbit. Coal, All Furn. Spreader-Stoker Fired

Data Source — Mean (x̄), Sample Standard Deviation (S), and Max. and Min. Values (lb of SO_2/ton) — **Sample Number**

20S[a] 30S 40S 50S 60S

Data Source	Sample Number
AP-42 '73 '82	
Ref. 4 (Bit. Coal)	49
Ref. 4 (Subbit. Coal)	15

Ref. 4 (Bit. Coal): 20S, 31.4S, 35S, 38S, 39.5S, 47.2S, 59.5S

Ref. 4 (Subbit. Coal): 22S, 26.3S, 35.0S, 43.7S, 51S

Data Source — Coefficient of Variation (CV), % and Relative Extreme Value (REV), %

−100 −50 0 50 100

Ref. 4 (Bit. Coal): −49, 20, 50

Ref. 4 (Subbit. Coal): −38, 25, 46

Comments: [a] Sulfur content in weight %

TABLE D.6 Completed Data Base Form for Emission Factors: SO_x Data, Form 4

Min. 1.0S 1.0S Max. ($\bar{x}$) — Emissions-Estimation Parameter Variability Data Base — Low REV 1.0CV 1.0CV High REV (0)

Date: 5/2/86

Parameter		Emission Factor, Species: SOx
SCC	Number	1-02-002-06, 1-03-002-07
	Description	Ext. Comb. Boiler, Ind., Comm-Inst., Bit. Coal, Underfed Stoker

Data Source	Mean ($\bar{x}$), Sample Standard Deviation (S), and Max. and Min. Values (lb of SO_2/ton)	Sample Number
AP-42 82	20S[a] 30S 40S 50S 60S; 31S	
Ref. 4	22.4S, 39.6S; 19.3S, 31S, 47.5S	9

Data Source	Coefficient of Variation (CV), % and Relative Extreme Value (REV), %
Ref. 4	−100 −50 0 50 100; −28, 28; −38, 53

Comments: a : Sulfur content in weight %.

TABLE D.7 Completed Data Base Form for Emission Factors: SO_x Data, Form 5

Min. 1.0S 1.0S Max. ($\bar{x}$)

Emissions-Estimation Parameter Variability Data Base

Low REV 1.0CV 1.0CV High REV (0)

Date: 4/14/86

Parameter		Emission Factor, Species: SOx
SCC	Number	1-01-002-:01, 02
	Description	Ext. Comb. Boiler; Elec. Gen., Bit. Coal, Pulv., Dry Bottom; Wet Bottom

Data Source	Mean ($\bar{x}$), Sample Standard Deviation (S), and Max. and Min. Values (lb of SO_2/ton)	Sample Number
	30S[a] 40S 50S 60S 70S	
AP-42 73 82	38.5 39.5	
Ref. 5 Vertical	36.8S 37.1S 43.5S 50.2S 54.5	5
Ref. 5 Front Wall	35.4S 43.3S 56S 68.7S 70.2S	5
Ref. 6 Corner	40.9S 41.9S 43.2S 45.9S	5
Ref. 6 Horiz.	37.6S 37.9S 41.9S 42.6S	4

Data Source	Coefficient of Variation (CV), % and Relative Extreme Value (REV), %
	100 −50 0 50 100
Ref. 5 Vertical	−15 15 24
Ref. 5 Front Wall	−37 23 25
Ref. 6 Corner	−7 5 5
Ref. 6 Horiz.	−6 5 7

Comments: a Sulfur content in weight %.

TABLE D.8 Completed Data Base Form for Emission Factors: SO_x Data, Form 6

Min. 1.0S 1.0S Max. — $\bar{x}$

Emissions-Estimation Parameter Variability Data Base

Low REV 1.0CV 1.0CV High REV — 0

Date: 4/14/86

Parameter		Emission Factor	Species: SOx
SCC	Number	1-01-004-01; 1-02-004-01; 1-03-004-01	
	Description	Ext. Comb. Boiler; Elec. Gen., Ind., Comm.-Inst.; Residual Oil	

Data Source	Mean ($\bar{X}$), Sample Standard Deviation (S), and Max. and Min. Values (lb of SO_2/ton)	Sample Number
	0S[a] 50S 100S 150S 200S	
AP-42 (73 EF 82	157 1575	
Ref. 7	107S 169S 20S 138S 162S	97

Data Source	Coefficient of Variation (CV), % and Relative Extreme Value (REV), %	
	−100 −50 0 50 100	
Ref. 7	23 −86 17	

Comments: a S is the sulfur content in weight %

TABLE D.9 Completed Data Base Form for Emission Factors: NO_x Data, Form 1

Min. 1.0S 1.0S Max. (x̄) — Emissions-Estimation Parameter Variability Data Base — Low REV 1.0CV 1.0CV High REV (0)

Date: 4/10/86

Parameter		Emission Factor	Species: NO_x
SCC	Number	1-01-002-01	
	Description	Ext. Comb. Boiler, Elec. Gen. Bit. Coal, Pulv., Wet Bottom	

Mean (x̄), Sample Standard Deviation (S), and Max. and Min. Values (lb of NO_2/ton)

Data Source	10	20	30	40	50	Sample Number
AP-42 ('73 EF ('82			30 34			
Ref. 1		16.1				1
Ref. 8 [a]			31.8 31.4	39.5	47.1 47.5	4
Ref. 1 & 8 [a]		16.1 22.6	34.8		47.0 47.5	5

Coefficient of Variation (CV), % and Relative Extreme Value (REV), %

Data Source	-100	-50	0	50	100
Ref. 1					
Ref. 8		-21	19 20		
Ref 1 & 8		-54	35 37		

Comments: [a] Conversion of data from lb/10^6 Btu to lb/ton is based on a coal heating value of 12,000 Btu/lb.

TABLE D.10 Completed Data Base Form for Emission Factors: NO_x Data, Form 2

Min. 1.0S 1.0S Max. ($\bar{x}$) — Emissions-Estimation Parameter Variability Data Base — Low REV 1.0CV 1.0CV High REV (0)

Date: 4/10/86

Parameter		Emission Factor	Species: NO_x
SCC	Number	1-01-002-02	
	Description	Ext. Comb. Boiler, Elec. Gen., Bit. Coal, Pulv., Dry Bottom	

Data Source — Mean ($\bar{x}$), Sample Standard Deviation (S), and Max. and Min. Values (lb of NO_2/ton) — Sample Number

Data Source	Mean ($\bar{x}$), Sample Standard Deviation (S), and Max. and Min. Values (lb of NO_2/ton)	Sample Number
AP-42 {'73 / EF {'82	18 (a); 15 (a); 21 (a)	
Ref. 1	11.2; 22.2; 11.0; 16.7; 21.6	4
Ref. 8A [b,c]	17.6; 25.4; 18.6; 21.5; 26.9	4
Ref. 8B [b,d]	13.3; 28.8; 11.8; 21.0; 31.2	8
Ref. 188	13.6; 26.4; 11.0; 20.0; 31.2	16

Data Source — Coefficient of Variation (CV), % and Relative Extreme Value (REV), %

Data Source	Coefficient of Variation (CV), % and Relative Extreme Value (REV), %
Ref. 1	33; −34; 30
Ref. 8A	18; −13; 25
Ref. 8B	37; −44; 48
Ref. 188	32; −45; 56

Comments:
a. For tangentially fired boilers
b. Conversion of data from lb/10^6 Btu to lb/ton is based on a coal heating value of 12,000 Btu/lb.
c. For single wall fired boilers.
d. For horizontally opposed fired boilers.

TABLE D.11 Completed Data Base Form for Emission Factors: NO_x Data, Form 3

Min. 1.0S $\bar{x}$ 1.0S Max. — Emissions-Estimation Parameter Variability Data Base — Low REV 1.0CV 0 1.0CV High REV

Date: 4/10/86

Parameter		Emission Factor — Species: NOx
SCC	Number	1-01-002-03
	Description	Ext. Comb. Boiler, Elec. Gen., Bit. Coal, Cyclone

Data Source	Mean ($\bar{x}$), Sample Standard Deviation (S), and Max. and Min. Values (lb of NO_2/ton)	Sample Number
	20 30 40 50 60	
AP-42 ('73 EF '82	37 (mean); 55 (max)	
Ref. 1	61.6	1
Ref. 8	28.8; 35.6; 43.1; 49	6
Ref. 1 & 8	27.5; 28.8; 39.3; 51.1; 61.6	7

Data Source	Coefficient of Variation (CV), % and Relative Extreme Value (REV), %
	−100 −50 0 50 100
Ref. 1	0
Ref. 8	−19; 21; 37
Ref. 1 & 8	−25; 30; 38

Comments:

TABLE D.12 Completed Data Base Form for Emission Factors: NO_x Data, Form 4

Min. 1.0S 1.0S Max. (x̄) — **Emissions-Estimation Parameter Variability Data Base** — Low REV 1.0CV 1.0CV High REV (0)

Date: 4/11/86

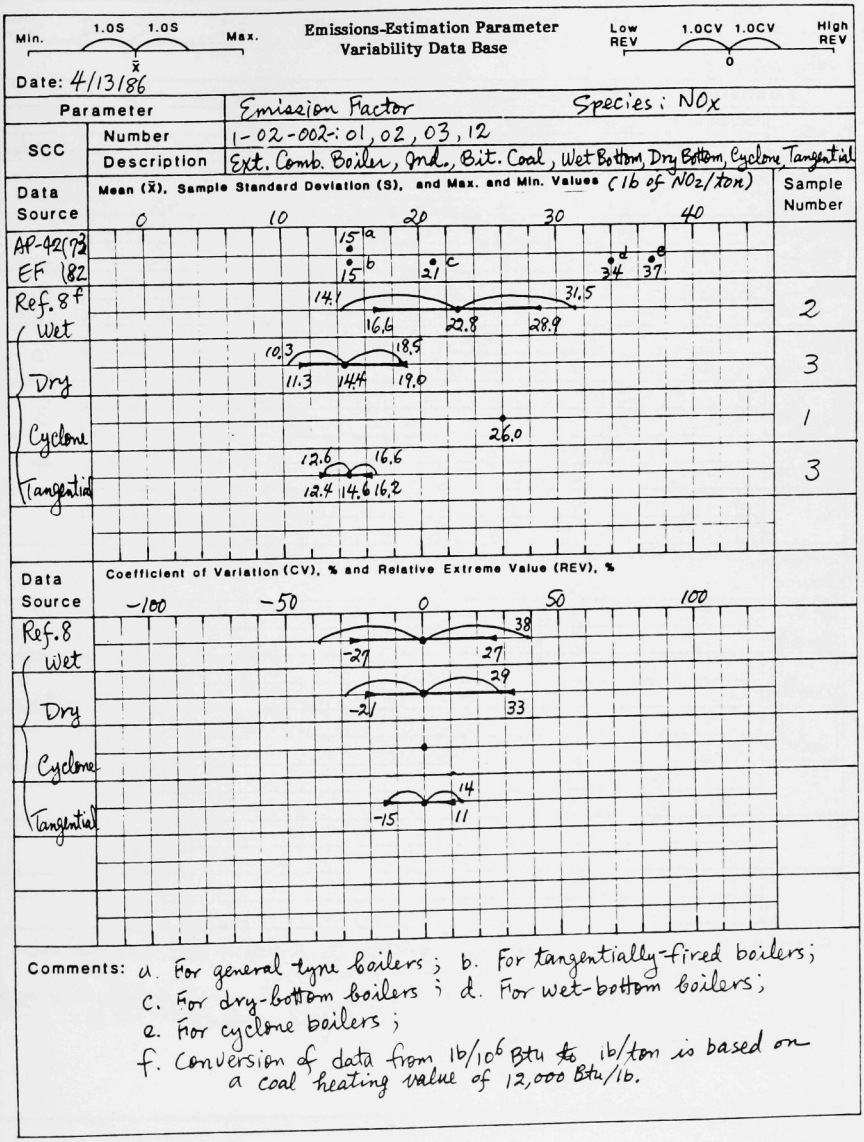
Parameter		Emission Factor	Species: NOx
SCC	Number	1-01-002-12	
	Description	Ext. Comb. Boiler, Elec. Gen., Bit. Coal., Pulv., Dry, Tangential	

Data Source	Mean (x̄), Sample Standard Deviation (S), and Max. and Min. Values (lb of NO_2/ton)	Sample Number
AP-42 EF, 82	15	
Ref. 8	13.0, 15.8; 11.8, 14.4, 17.3	17

Data Source	Coefficient of Variation (CV), % and Relative Extreme Value (REV), %
Ref. 8	10; −19, 19

Comments:

TABLE D.13 Completed Data Base Form for Emission Factors: NO_x Data, Form 5



Comments: a. For general type boilers; b. For tangentially-fired boilers; c. For dry-bottom boilers; d. For wet-bottom boilers; e. For cyclone boilers; f. Conversion of data from lb/10^6 Btu to lb/ton is based on a coal heating value of 12,000 Btu/lb.

TABLE D.14 Completed Data Base Form for Emission Factors: NO_x Data, Form 6

Min. 1.0S 1.0S Max. ($\bar{x}$) — Emissions-Estimation Parameter Variability Data Base — Low REV 1.0CV 1.0CV High REV (0)

Date: 4/13/86

Parameter		Emission Factor	Species: NO_x
SCC	Number	1-02-002	
	Description	Ext. Comb. Boiler; Gnd., Bit. Coal; Spreader, Overfeed, Underfeed Stoker	

Data Source	Mean ($\bar{x}$), Sample Standard Deviation (S), and Max. and Min. Values (lb of NO_2/ton)[a] (0 10 20 30 40)	Sample Number
AP-42 {'73 EF '82	6[b]; 15[c]; 7.5[d] 9.5[e] 14[f]	
Ref 1 [a] Spreadr	0.4	1
Ref 8 [a] Spreader	12.5 17.9; 10.4 15.2 18.9	10
Overfeed	5.5 9.5; 5.9 6.5 9.5	3
Underfeed	8.4 10.8; 8.6 9.6 11.2	4

Data Source	Coefficient of Variation (CV), % and Relative Extreme Value (REV), % (−50 0 50)
Ref. 8 Spreader	-31; 18 25
Overfeed	-11; 14 16
Underfeed	10; 13 17

Comments: a Conversion of data from lb/10^6 Btu to lb/ton of coal is based on a coal heating value of 12,000 Btu/lb;
b For spreader stoker < 10^6 mm Btu/hr.; c For spreader stoker 10-100 mm Btu/hr
d. For overfeed stoker; e For Underfeed stoker;
f. For spreader stoker.

TABLE D.15 Completed Data Base Form for Emission Factors: NO_x Data, Form 7

Min. 1.0S 1.0S Max. — $\bar{x}$

Emissions-Estimation Parameter Variability Data Base

Low REV 1.0CV 1.0CV High REV — 0

Date: 5/5/86

Parameter		Emission Factor — Species NO_x
SCC	Number	1-01-002-01,02, 21,22; 1-02-002-01, 02, 21, 22; 1-03-002-05, 06, 21, 22
	Description	Ext. Comb. Boiler, Elec. Gen., Ind., Comm.-Inst., Bit. & Subbit. Coal, Wet & Dry Bottom

Data Source	Mean ($\bar{x}$), Sample Standard Deviation (S), and Max. and Min. Values (lb of NO_2/ton)	Sample Number
AP-42: 82	a: 21; b: 34	
Ref. 4 (Dry Bottom)	9.5; 12.91; 21.1; 28.3; 44.5	28
(Wet Bottom)	9.8; 16.5; 33.8; 50.5; 57.8	2

Data Source	Coefficient of Variation (CV), % and Relative Extreme Value (REV), %
Ref. 4 (Dry Bottom)	-55; 39; 111
(Wet Bottom)	-57; 49; 71

Comments:

a. Dry bottom;

b. Wet bottom

TABLE D.16 Completed Data Base Form for Emission Factors: NO_x Data, Form 8

Min. 1.0S 1.0S Max. (x̄) — Emissions-Estimation Parameter Variability Data Base — Low REV 1.0CV 1.0CV High REV (0)

Date: 5/5/86

Parameter		Emission Factor — Species NOx
SCC	Number	1-01-002-03, 23; 1-02-002-03, 23
	Description	Ext. Comb. Boiler, Elec. Gen. & Ind., Bit. & Subbit. Coal, Cyclone

Data Source	Mean (x̄), Sample Standard Deviation (S), and Max. and Min. Values (lb of NO_2/ton)	Sample Number
AP-42; 2C	10 20 30 40 50; 39.5	
Ref. 4	27.7 (27.2) 45.3 (36.5) 54.5	7

Data Source	Coefficient of Variation (CV), % and Relative Extreme Value (REV), %
Ref. 4	−100 −50 0 50 100; −26 24 49

Comments:

TABLE D.17 Completed Data Base Form for Emission Factors: NO_x Data, Form 9

Min. 1.0S 1.0S Max. (x̄) — Emissions-Estimation Parameter Variability Data Base — Low REV 1.0CV 1.0CV High REV (0)

Date: 5/5/86

Parameter		Emission Factor — Species NO_x
SCC	Number	1-01-002-; 12, 26; 1-02-002-: 12, 26; 1-03-002-: 16, 26
	Description	Ext. Comb. Boiler, Elec. Gen., Ind., Comm.-Inst., Bit. & Subbit. Coal, Tangential

Data Source	Mean (x̄), Sample Standard Deviation (S), and Max. and Min. Values (lb of NO_2/ton)	Sample Number
AP-42:82	15	
Ref. 4	10.5, 10.9, 14.7, 19, 23.1	29

Data Source	Coefficient of Variation (CV), % and Relative Extreme Value (REV), %
Ref. 4	−29, 39, 57

Comments:

TABLE D.18 Completed Data Base Form for Emission Factors: NO_x Data, Form 10

Min. 1.0S 1.0S Max. (x̄)

Emissions-Estimation Parameter Variability Data Base

Low REV 1.0CV 1.0CV High REV (0)

Date: 5/5/86

Parameter		Emission Factor — Species: NO_x
SCC	Number	1-01-002-04, 24 ; 1-02
	Description	Ext. Comb. Boiler, Elec. Gen., Ind., Comm. Inst., Bit. & Subbit Coal, Spreader Stoker

Data Source	Mean (x̄), Sample Standard Deviation (S), and Max. and Min. Values (16 of NO_2/ton)	Sample Number
	0 10 20 30 40	
AP-42, '82	14	
Ref. 4	10.9 (13.7), 16.5 (17.5), 7.0	35

Data Source	Coefficient of Variation (CV), % and Relative Extreme Value (REV), %	
	−100 −50 0 50 100	
Ref. 4	20, 28, −49	

Comments:

TABLE D.19 Completed Data Base Form for Emission Factors: NO_x Data, Form 11

Min. 1.0S 1.0S Max. — $\bar{x}$

Emissions-Estimation Parameter Variability Data Base

Low REV 1.0CV 1.0CV High REV — 0

Date: 5/5/86

Parameter		Emission Factor Species NOx
SCC	Number	1-01-002-05,25; 1-02-002-05,10,25; 1-03-002-07,11,25
	Description	Ext. Comb. Boiler, Elec. Gen., Ind., Comm.-Inst., Bit. & Subbit. Coal, Overfeed Stoker

Data Source	Mean ($\bar{x}$), Sample Standard Deviation (S), and Max. and Min. Values (lb of NO_2/ton)	Sample Number
	0 5 10 15 20	
AP-42:82	7.5	
Ref. 4	5.8 9.0 — 5.0 7.4 9.5	15

Data Source	Coefficient of Variation (CV), % and Relative Extreme Value (REV), %
	-100 -50 0 50 100
Ref. 4	21 — -32 28

Comments:

TABLE D.20 Completed Data Base Form for Emission Factors: NO_x Data, Form 12

Min. 1.0S 1.0S Max. ($\bar{x}$)

Emissions-Estimation Parameter Variability Data Base

Low REV 1.0CV 1.0CV High REV (0)

Date: 5/5/86

Parameter		Emission Factor — Species: NOx
SCC	Number	1-02-002-06; 1-03-002-07
	Description	Ext. Comb. Boiler, Ind., Comm.-Inst., Bit. Coal, Underfeed Stoker

Data Source	Mean ($\bar{x}$), Sample Standard Deviation (S), and Max. and Min. Values (lb of NO_2/ton)	Sample Number
AP-42, 82	0 5 10 15 20; 9.5; 7.2, 11.8; 7.0, 9.5, 13.5	8
Ref. 4		

Data Source	Coefficient of Variation (CV), % and Relative Extreme Value (REV), %
Ref. 4	−100 −50 0 50 100; −26, 23, 42

Comments:

TABLE D.21 Completed Data Base Form for Emission Factors: NO_x Data, Form 13

Min. 1.0S 1.0S Max. ($\bar{x}$)

Emissions-Estimation Parameter Variability Data Base

Low REV 1.0CV 1.0CV High REV (0)

Date: 4/14/86

Parameter		Emission Factor — Species: NO_x
SCC	Number	1-01-002-: 01, 02
	Description	Ext. comb. Boiler; Elec. Gen.; Bit. Coal, Pulv.; Dry Bottom & Wet Bottom

Data Source	Mean ($\bar{x}$), Sample Standard Deviation (S), and Max. and Min. Values (lb of NO_2/ton)	Sample Number
AP-42 (73 EF 82)		
Ref. 5 (Vertical	5.2, 6.6, 10.8, 15.0, 16.7	5
Front Wall)	17.2, 18.9, 23.9, 27.8, 29.0	5
Ref. 6 (Corner	11.9, 12.9, 15.2, 17.4, 17.2	4
Horiz.	13.2, 13.0, 15.4, 17.8, 18.7	4

Data Source	Coefficient of Variation (CV), % and Relative Extreme Value (REV), %
Ref. 5 (Vertical	-52, 39, 55
Front Wall)	-26, 19, 24
Ref. 6 (Corner	-21, 15, 14
Horiz.	-14, 15, 22

Comments:

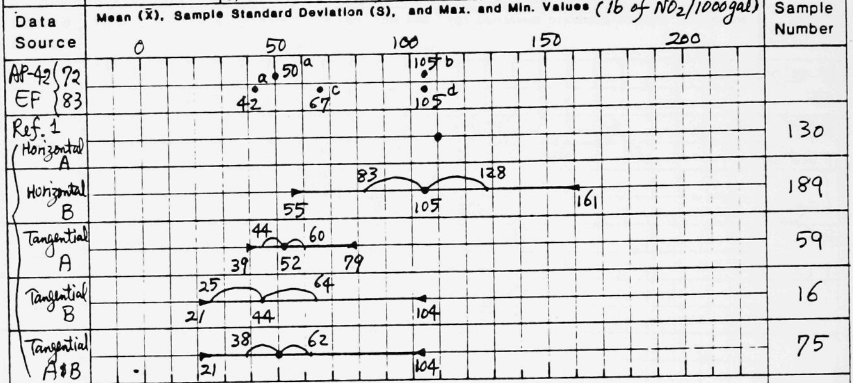
TABLE D.22 Completed Data Base Form for Emission Factors: NO_x Data, Form 14

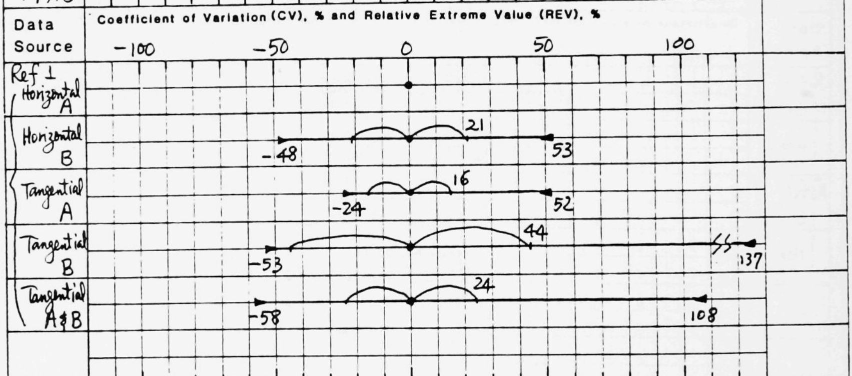
Emissions-Estimation Parameter Variability Data Base

Min. 1.0S 1.0S Max. ($\bar{x}$) — Low REV 1.0CV 1.0CV High REV (0)

Date: 4/13/86

Parameter		Emission Factor Species : NOx
SCC	Number	1-01-004-:01, 05 ; 04, 06
	Description	Ext. Comb. Boiler, Elec. Gen., Residual Oil. No. 6 & 5, Normal & Tangential





Comments:
a For tangentially-fired boilers
b For general type boilers
c For boilers other than vertically-or tangentially-fired
d For vertically-fired boilers
e Data distribution is not available.

TABLE D.23 Completed Data Base Form for Emission Factors: NO_x Data, Form 15

Min. 1.0S 1.0S Max. — $\bar{x}$

Emissions-Estimation Parameter Variability Data Base

Low REV 1.0CV 1.0CV High REV — 0

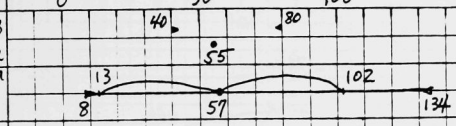
Date: 4/14/86

Parameter		Emission Factor	Species: NOx
SCC	Number	1-02-004-: 01, 02, 03, 04; 1-03-004-: 01, 02, 03, 04	
	Description	Ext. Comb. Boiler, Ind., Comm.-Inst., Residual Oil #6, #5	

Data Source — Mean ($\bar{x}$), Sample Standard Deviation (S), and Max. and Min. Values (lb of NO_2/1000 gal) — Sample Number

AP-42 {73, 82} EF

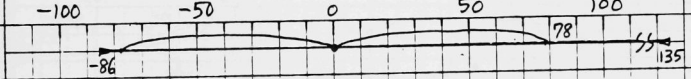
Ref. 1[a]



17

Data Source — Coefficient of Variation (CV), % and Relative Extreme Value (REV), %

Ref. 1



Comments: a Only individual data, not mean values, are used.

TABLE D.24 Completed Data Base Form for Emission Factors: NO_x Data, Form 16

Min. 1.0S 1.0S Max. (x̄) — Emissions-Estimation Parameter Variability Data Base — Low REV 1.0CV 1.0CV High REV (0)

Date: 4/14/86

Parameter		Emission Factor	Species: NOx
SCC	Number	1-01-006-01, 04	
	Description	Ext. Comb. Boiler; Elec. Gen. (>100 MM Btu/hr; Tangential), Natural Gas	

Data Source	Mean (x̄), Sample Standard Deviation (S), and Max. and Min. Values (lb of $NO_2/10^6$ ft^3)	Sample Number
	100 250 500 750 1,000 1200	
AP-42 {73 EF {82	275[a] 600 550	
Ref 1[b]	116 100 428 740 1360	27
Ref 1A[c]	184 100 576 968 1360	13
Ref 1B[d]	187 111 290 393 444	14

Data Source	Coefficient of Variation (CV), % and Relative Extreme Value (REV), %
	−100 −50 0 50 100
Ref 1[b]	−77 73 218
Ref 1A[c]	−83 68 136
Ref 1B[d]	−62 35 53

Comments:
a. For tangentially fired boilers
b. Original data set is from 4 sources. Used data from 3 sources neglecting one with one typical data.
c. Data for all boilers except for tangentially fired boilers
d. Tangentially fired boilers

TABLE D.25 Completed Data Base Form for Emission Factors: NO_x Data, Form 17

Min. 1.0S 1.0S Max. (x̄) — Emissions-Estimation Parameter Variability Data Base — Low REV 1.0CV 1.0CV High REV (0)

Date: 4/14/86

Parameter		Emission Factor — Species: NO_x
SCC	Number	1-02-006-02
	Description	Ext. Comb. Boiler, Ind., Natural Gas (10 – 100 MM Btu/hr)

Data Source	Mean (x̄), Sample Standard Deviation (S), and Max. and Min. Values (lb of NO_2/10^6 ft^3)	Sample Number
AP-42 (73 EF 182	120 — 140 — 230	
Ref. 1	64 — 142 — 296 — 528 — 620	4

Data Source	Coefficient of Variation (CV), % and Relative Extreme Value (REV), %
Ref. 1	−100 … 50 … 0 … 50 … 100; −52 … 74 … 189

Comments:

TABLE D.26 Completed Data Base Form for Emission Factors: NO_x Data, Form 18

Min. 1.0S 1.0S Max. ($\bar{x}$) — Emissions-Estimation Parameter Variability Data Base — Low REV 1.0CV 1.0CV High REV (0)

Date: 4/14/86

Parameter		Emission Factor	Species: NOx
SCC	Number	1-03-006-03	
	Description	Ext. Comb. Boiler, Comm.-Inst., and Domestic (<10 MM Btu/hr) Natural Gas	

Data Source: Mean ($\bar{x}$), Sample Standard Deviation (S), and Max. and Min. Values (lb of $NO_2/10^6 ft^3$) — Sample Number

Data Source	0	100	200	300	400	Sample Number
AP-42 (73 / 82 EF)		80 — 100 — 120				
Ref. 1	-2 — 27 — 90 — 183 — 354					14

Data Source: Coefficient of Variation (CV), % and Relative Extreme Value (REV), %

Data Source	-100	-50	0	50	100
Ref. 1	-103 — -59 — 103 — 55 — 292				

Comments: